Bayesian Regularized U-MIDAS under Extreme Frequency Mismatch

Xiaoyan Zhou
Beijing Normal–Hong Kong Baptist University

Joshua Chan
Purdue University

Dan Zhu
Monash University

April 2026

Abstract

Unrestricted MIDAS (U-MIDAS) estimates a coefficient for each high-frequency lag and often outperforms parametric MIDAS when the frequency mismatch is modest. However, it becomes unstable when the number of high-frequency bins m is large relative to the low-frequency sample size T. Using a continuous-time embedding, we show that under extreme frequency mismatch, unrestricted MIDAS attempts to recover more information than the data can support, leading to an ill-posed inverse problem for the lag-weight function. We establish that the unregularized estimator is consistent when $m = o(T/\log T)$ and that economically relevant fixed linear functionals are asymptotically normal. To address more severe mismatch, we introduce smoothing-spline regularization that controls the effective degrees of freedom and delivers valid inference for linear functionals even when m grows arbitrarily fast. A Bayesian implementation yields asymptotically valid credible intervals. Simulations and two empirical applications, US GDP nowcasting and macroeconomic uncertainty, demonstrate the stability and relevance of the approach. Overall, our results justify inference on economically interpretable within-period timing measures even when pointwise lag estimates are unreliable.

Keywords: mixed-frequency data; smoothing spline regularization; linear functionals; Bayesian inference; nowcasting; macroeconomic uncertainty.
JEL classification: C13; C22; C32; C53.

1 Introduction

High-frequency macroeconomic and financial indicators—such as card spending measures, interest rates, credit spreads, and indicators of financial stress—respond quickly to news and changing expectations, making them natural inputs for real-time monitoring and forecasting of lower-frequency macroeconomic outcomes (see, e.g., Giannone et al., 2008; Altissimo et al., 2010; Doz et al., 2012). In applied work, such variables are widely used to reduce information delays and improve nowcasting and forecasting performance. Mixed-data sampling (MIDAS) regressions provide a parsimonious framework for combining predictors observed at different frequencies (Ghysels et al., 2004, 2007; Andreou et al., 2010). Standard (restricted) MIDAS models impose a low-dimensional parametric structure on the lag weights, which can stabilize estimation but may be restrictive when the lag profile is irregular. Unrestricted MIDAS (U-MIDAS) relaxes these restrictions by estimating a coefficient for each high-frequency lag (or bin) (Foroni et al., 2015), and is often empirically competitive with parametric MIDAS when the frequency mismatch is modest.

A central challenge arises when the frequency mismatch is large. In a U-MIDAS regression, the number of high-frequency lags per low-frequency period, denoted by m, can be comparable to or even exceed the number of low-frequency observations T, particularly when daily or intraday data are used to forecast quarterly macro series. In such cases, U-MIDAS can become numerically and statistically unstable: regressors are highly collinear, coefficient estimates are noisy, and standard large-sample approximations may fail. This instability reflects a deeper identification problem that arises when the lag profile is treated as an unrestricted high-dimensional object.

This paper provides a theoretical and practical resolution to that problem. Our starting point is to embed the U-MIDAS regression in a continuous-time framework. When the high-frequency regressor is viewed as a discretely observed diffusion process, the U-MIDAS regression on high-frequency lags can be interpreted as a Riemann-Itô approximation to a stochastic integral with an unknown lag-weight function. This perspective makes clear why extreme mismatch is difficult: as m grows, estimating an unrestricted lag-weight function from T low-frequency observations becomes an ill-posed inverse problem. This insight motivates regularization schemes that impose structure on the lag profile while preserving the flexibility of U-MIDAS.

We first establish asymptotic results for U-MIDAS under joint growth of T and the frequency mismatch m, with a focus on the low-dimensional summaries that are central to empirical MIDAS analysis. These summaries—such as cumulative effects of high-frequency

variables over the within-period horizon and measures of early- versus late-period information content—are naturally expressed as linear functionals of the lag-weight function and capture how much information arrives within a period and when it matters most. Working in a Hilbert space of square-integrable lag-weight functions, we show that the unregularized U-MIDAS estimator is consistent under the moderate-growth condition $m = o(T/\log T)$ and that fixed linear functionals of the lag-weight function are asymptotically normal. This clarifies when unrestricted estimation remains valid despite an expanding lag dimension and provides a benchmark for understanding instability under more severe mismatch.

We then introduce smoothing-spline regularization that stabilizes U-MIDAS estimation under more extreme frequency mismatch. The penalty controls the roughness of the lag-weight function and reduces the effective complexity of the problem through its degrees of freedom. With an appropriately scaled penalty, we establish consistency and asymptotic normality for linear functionals of the lag-weight function even when the number of high-frequency bins m grows arbitrarily fast, provided the effective degrees of freedom increase sufficiently slowly. This reframes the relevant asymptotic condition from the nominal dimension m to an operational notion of model complexity and delivers valid inference for economically interpretable within-period timing summaries.

Finally, we implement the regularization in a Bayesian framework, in the spirit of the Bayesian nonparametric approaches developed by Koop and Poirier (2004) and Koop et al. (2005), and show that credible intervals for linear functionals are asymptotically valid under standard Bernstein–von Mises arguments (Castillo and Nickl, 2014).

We illustrate the practical value of the approach in two applications. The first concerns nowcasting US quarterly GDP growth using a rich set of monthly, weekly, and daily predictors. The second examines the role of high-frequency financial information in explaining macroeconomic uncertainty measured by disagreement in the Survey of Professional Forecasters, connecting to the literature on uncertainty measurement and forecast-based indices (Jurado et al., 2015; Rossi and Sekhposyan, 2015). Across both applications, regularized U-MIDAS substantially improves stability relative to unregularized U-MIDAS and delivers predictive performance comparable to or better than a LASSO-type benchmark.

Our paper contributes to a growing literature on shrinkage and regularization for U-MIDAS and related mixed-frequency regressions. Existing work emphasizes sparsity and variable selection in high-dimensional settings (Mogliani and Simoni, 2021; Babii et al., 2022) or develops inference under joint ("double") asymptotics for within-period and across-period sample sizes (Ghosh and Linton, 2023). More broadly, our analysis is also related to the literature on nonparametric series estimation and penalized regression. Classical results

show that sieve and series estimators can deliver consistency and asymptotic normality for low-dimensional functionals under increasing dimension, even when the underlying object of interest is infinite-dimensional (Newey, 1997; Chen and Christensen, 2015). Our spline-based regularization is similarly connected to the penalized-spline literature, where roughness penalties control effective model complexity and stabilize large-sample behavior (Claeskens et al., 2009).

What is specific to our setting, however, is the mixed-frequency structure. By embedding U-MIDAS in continuous time, we show that under extreme frequency mismatch the instability of unrestricted U-MIDAS can be understood as an ill-posed inverse problem for the lag-weight function. This perspective motivates smoothing-spline regularization, which controls the effective degrees of freedom and delivers consistent estimation and valid inference for economically meaningful linear functionals rather than pointwise lag coefficients. Our discussion is also related in spirit to recent work on ridgeless and near-ridgeless high-dimensional least squares, which highlights how regularization affects behavior when dimensionality is large relative to sample size (Hastie et al., 2022).

The remainder of the paper is organized as follows. Section 2 develops the continuous-time embedding and asymptotic theory for unregularized and spline-regularized U-MIDAS, including inference on linear functionals. Section 3 presents the Bayesian implementation and establishes the asymptotic validity of credible intervals. Section 4 reports Monte Carlo evidence on instability under extreme mismatch and the stabilizing role of regularization. Section 5 applies the method to US GDP nowcasting, and Section 6 studies macroeconomic uncertainty using Survey of Professional Forecasters disagreement.

2 U-MIDAS under Extreme Frequency Mismatch

This section develops asymptotic theory for U-MIDAS under increasing frequency mismatch. The key idea is to view the high-frequency regressor as a discretely observed continuous-time process and to interpret the U-MIDAS regression as a Riemann–Itô approximation to a stochastic integral. This embedding provides a transparent large-sample characterization of U-MIDAS when the number of within-period lags m grows with the low-frequency sample size T, and clarifies why unrestricted estimation becomes unstable under extreme mismatch.

2.1 Continuous-Time Embedding and High-Dimensional U-MIDAS

We consider a J-dimensional high-frequency regressor process $\boldsymbol{X}(s)$ observed on a fine grid within each low-frequency period. Let $\mathcal{F}_t^X$ denote the filtration generated by $\{\boldsymbol{X}(r) : r \le t\}$, and let $\mathcal{F}_{t^-}^X$ denote its left limit. We assume throughout that the lag-weight function satisfies $\boldsymbol{\beta} \in H := L^2([0,\tau];\mathbb{R}^J)$, where τ denotes the within-period horizon (e.g., one quarter). Under the diffusion specification introduced below, this is sufficient for the stochastic integral

$$\int_0^\tau \boldsymbol{\beta}(s)^\top d\boldsymbol{X}(t-s)$$

to be well defined in the Itô sense.

The low-frequency outcome $\{y_t\}$ depends on past high-frequency increments through the unknown lag-weight function $\boldsymbol{\beta}(\cdot)$. In continuous time, we write

$$y_t = \alpha + \int_0^\tau \boldsymbol{\beta}(s)^\top \, d\boldsymbol{X}(t-s) + u_t, \qquad t = 1, 2, \ldots,$$

where u_t is a low-frequency disturbance satisfying $\mathbb{E}[u_t \mid \mathcal{F}_{t^-}^X] = 0$. The function $\boldsymbol{\beta}(s)$ summarizes how shocks occurring at different points within the period contribute to y_t, and therefore provides a direct representation of within-period timing effects.

To connect this representation to U-MIDAS, partition $[0,\tau]$ into m equal bins of width $h = \tau/m$, and define the high-frequency increments

$$\Delta \boldsymbol{X}_{t,\ell} = \boldsymbol{X}\big(t - (\ell-1)h\big) - \boldsymbol{X}\big(t - \ell h\big), \qquad \ell = 1, \ldots, m.$$

The discrete U-MIDAS regression on binned high-frequency data can then be written as

$$y_t = \alpha + \sum_{\ell=1}^{m} \boldsymbol{\beta}\left((\ell-1)h\right)^\top \Delta \boldsymbol{X}_{t,\ell} + u_t, \qquad t = 1, 2, \ldots,$$

which is a Riemann-Itô approximation to the stochastic integral above. Since step functions are dense in $H = L^2([0,\tau];\mathbb{R}^J)$, the discretized m-bin specification provides a natural sieve approximation to the continuous-time lag-weight function as $m \to \infty$. When m is large, this discretization yields a high-dimensional regression with a coefficient vector of dimension mJ. Under extreme frequency mismatch, unrestricted estimation becomes unstable because the mapping from the lag-weight function $\boldsymbol{\beta}(\cdot)$ to the observables $\{y_t\}$ is a compact operator. Recovering $\boldsymbol{\beta}(\cdot)$ from finitely many low-frequency observations is therefore ill-posed without additional structure, motivating the regularization approach developed below.

We adopt the following baseline assumption on the high-frequency regressor process.

Assumption 1. *Let* $\boldsymbol{X}(s) = \big(X_1(s), \ldots, X_J(s)\big)^\top$ *denote a* J*-dimensional high-frequency regressor process satisfying*

$$d\boldsymbol{X}(s) = \boldsymbol{\mu}\, ds + \boldsymbol{\Sigma}\, d\boldsymbol{W}(s),$$

where $\boldsymbol{W}(s)$ *is a* J*-dimensional standard Brownian motion,* $\boldsymbol{\mu} \in \mathbb{R}^J$ *is a fixed drift vector, and* $\boldsymbol{\Sigma} \in \mathbb{R}^{J\times J}$ *is a fixed volatility matrix. The quadratic covariation matrix* $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top$ *is finite and positive definite.*

Extensions to time-varying drift and volatility coefficients are possible but not pursued here. Since our objective is not to estimate $\boldsymbol{\mu}(\cdot)$ or $\boldsymbol{\Sigma}(\cdot)$, but rather to study the behavior of U-MIDAS estimators as functions of the lag-weight profile $\boldsymbol{\beta}(\cdot)$, we maintain constant coefficients for theoretical clarity.

Before proceeding to the formal asymptotic analysis, it is useful to clarify why our theoretical results focus on linear functionals of the lag-weight function rather than the full function itself. In MIDAS applications, the primary objects of economic interest are typically low-dimensional summaries of $\boldsymbol{\beta}(\cdot)$—such as cumulative within-period effects, early- versus late-period information content, or smoothly weighted timing measures—rather than pointwise values of the lag profile. These summaries can all be written as linear functionals of $\boldsymbol{\beta}(\cdot)$ and provide interpretable measures of how much high-frequency information is incorporated and when it matters most within a period. From a statistical perspective, such functionals are also the objects for which stable inference can be achieved under increasing frequency mismatch, even when estimation of the full lag-weight function becomes ill-posed. This motivates our emphasis on asymptotic results for linear functionals throughout the remainder of this section.

2.2 Hilbert-Space Formulation and Asymptotic Theory

To formalize the asymptotic analysis, recall that $H := L^2([0,\tau];\mathbb{R}^J)$ is the separable Hilbert space of square-integrable lag-weight functions, equipped with the inner product

$$\langle \boldsymbol{f}, \boldsymbol{g}\rangle_H = \int_0^\tau \boldsymbol{f}(s)^\top \boldsymbol{g}(s)\, ds, \qquad \boldsymbol{f}, \boldsymbol{g} \in H.$$

For each t and $\boldsymbol{f} \in H$, define the linear operator

$$Z_t(\boldsymbol{f}) := \int_0^\tau \boldsymbol{f}(s)^\top\, d\boldsymbol{X}(t-s),$$

which maps a lag-weight function to the corresponding aggregate of high-frequency innovations. With this notation, the continuous-time U-MIDAS regression can be written compactly as

$$y_t = \alpha + Z_t(\boldsymbol{\beta}) + u_t.$$

This formulation allows the collection of lag coefficients to be treated as elements of a function space, with inner products capturing dependence across lags. The discretized U-MIDAS estimator can then be viewed as a finite-dimensional approximation to an infinite-dimensional least-squares problem in H.

Let $H_m \subset H$ denote the subspace of J-valued step functions defined on the m equal-width bins

$$I_\ell = [(\ell - 1)\tau/m, \ell\tau/m), \qquad \ell = 1, \ldots, m.$$

Define the sample and population criteria

$$S_{T,m}(\boldsymbol{f}) := \frac{1}{T}\sum_{t=1}^{T} \big(y_t - Z_t(\boldsymbol{f})\big)^2, \qquad L(\boldsymbol{f}) := \mathbb{E}\big[(y_t - Z_t(\boldsymbol{f}))^2\big], \qquad \boldsymbol{f} \in H.$$

Ordinary least squares with m bins corresponds to minimizing $S_{T,m}$ over H_m. In function space, the estimator can be written as

$$\widehat{\boldsymbol{f}}_{T,m} := \Pi_m \widehat{\boldsymbol{\beta}}_{T,m} \in H_m,$$

where Π_m denotes the step-function embedding that maps the bin coefficients $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)'$ into H according to

$$(\Pi_m \boldsymbol{\theta})(s) := \sum_{\ell=1}^{m} \mathbf{1}_{[(\ell-1)\tau/m, \ell\tau/m)}(s)\, \theta_\ell, \qquad s \in [0, \tau].$$

The goal is to establish that the discrete criterion $S_{T,m}$ converges to its population counterpart L as $T, m \to \infty$, and that the corresponding minimizers $\widehat{\boldsymbol{f}}_{T,m}$ consistently recover the population-optimal lag-weight function in H. Moreover, we characterize the large-sample distribution of economically relevant linear functionals of the estimator. To state these results, it is convenient to introduce the population covariance and long-run variance operators that govern the asymptotic behavior of the U-MIDAS estimator.

Define $\boldsymbol{A} := \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top$. Let $Q : H \to H$ denote the population covariance operator,

$$(Q\boldsymbol{g})(s) = \boldsymbol{A}\,\boldsymbol{g}(s) + \boldsymbol{\mu}\boldsymbol{\mu}^\top \int_0^\tau \boldsymbol{g}(u)\,du, \qquad s \in [0,\tau].$$

Let $\Omega : H \to H$ denote the long-run covariance operator associated with $\{u_t Z_t(\cdot)\}$, defined by

$$\langle \boldsymbol{f}, \Omega \boldsymbol{g}\rangle_H = \Big(\sum_{k\in\mathbb{Z}} \gamma_u(k)\Big)\Big(\int_0^\tau \boldsymbol{f}(s)^\top \boldsymbol{\mu}\,ds\Big)\Big(\int_0^\tau \boldsymbol{g}(s)^\top \boldsymbol{\mu}\,ds\Big) + \sum_{|k|\le\tau} \gamma_u(k)\, I_k(\boldsymbol{f},\boldsymbol{g}),$$

where $\gamma_u(k) = \mathbb{E}[u_t u_{t-k}]$ and

$$I_k(\boldsymbol{f},\boldsymbol{g}) = \begin{cases} \displaystyle\int_k^\tau \boldsymbol{f}(s)^\top \boldsymbol{A}\,\boldsymbol{g}(s-k)\,ds, & k \ge 0, \\ \displaystyle\int_0^{\tau+k} \boldsymbol{f}(s)^\top \boldsymbol{A}\,\boldsymbol{g}(s-k)\,ds, & k < 0. \end{cases}$$

Theorem 1. *Let $\{u_t\}_{t\in\mathbb{Z}}$ be a strictly stationary process, independent of $\boldsymbol{W}$, satisfying $\mathbb{E}[u_t \mid \mathcal{F}_{t^-}^X] = 0$ and $\mathbb{E}|u_t|^{2+\delta} < \infty$ for some $\delta > 0$. Suppose further that $\{u_t\}$ is α-mixing with strong-mixing coefficients*

$$\alpha_u(k) := \sup_{A\in\sigma(u_t:t\le 0),\, B\in\sigma(u_t:t\ge k)} \big|\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)\big|, \quad k \ge 1,$$

and that $\sum_{k\ge1} \alpha_u(k)^{\delta/(2+\delta)} < \infty$. Let $m = m_T \to \infty$ with $m_T = o(T/\log T)$. Then:

1. *$\Pi_{m_T}\widehat{\boldsymbol{\beta}}_{T,m_T} \xrightarrow{p} \boldsymbol{\beta}$ in H.*
2. *For any fixed $\boldsymbol{\psi} \in H$,*

$$\sqrt{T}\,\big\langle \Pi_{m_T}\widehat{\boldsymbol{\beta}}_{T,m_T} - \boldsymbol{\beta}, \boldsymbol{\psi}\big\rangle_H \xrightarrow{d} \mathcal{N}\big(0, \big\langle \boldsymbol{\psi}, Q^{-1}\Omega Q^{-1}\boldsymbol{\psi}\big\rangle_H\big).$$

A proof of this theorem is provided in Appendix A. These results show that the unregularized U-MIDAS estimator remains consistent in Hilbert space as the frequency mismatch grows, provided the number of high-frequency bins satisfies $m = o(T/\log T)$. Moreover, for any fixed linear functional of the lag-weight function, the estimator is asymptotically normal, yielding valid large-sample inference for economically interpretable summaries of within-period effects. The operator representation of the asymptotic variance also clarifies efficiency properties relative to parametric MIDAS models and provides a rigorous benchmark for the regularization strategies introduced below.

2.3 Smoothing-spline Regularization

In many applications, the frequency mismatch can be sufficiently large that the moderate-growth condition $m = o(T/\log T)$ is restrictive. For example, using daily or intraday financial variables to forecast quarterly macroeconomic outcomes can yield hundreds (or more) high-frequency observations per low-frequency period, while the available low-frequency sample may span only a few decades. In such settings, unrestricted U-MIDAS estimation is not only high-dimensional but also economically implausible if it implies highly erratic within-period weights. In most macro-finance applications, the effect of high-frequency variables on low-frequency aggregates is expected to vary smoothly with timing within the period, reflecting gradual information aggregation, delayed transmission, and overlapping measurement windows. This motivates imposing smoothness directly on the lag-weight function and regularization schemes that remain well behaved even when m is large.

To relax the restriction $m = o(T/\log T)$, we introduce a smoothing-spline penalty on the lag-weight function $\boldsymbol{\beta}(\cdot)$. Specifically, let $P : H \to H$ be a self-adjoint, positive semi-definite penalty operator. For any $\boldsymbol{f} \in H$, the penalty takes the quadratic form $\langle \boldsymbol{f}, P\boldsymbol{f} \rangle_H$. Its finite-dimensional counterpart is a penalty matrix $\boldsymbol{A}_m$ that approximates P on the step-function subspace H_m. In particular, for any $\boldsymbol{f} = \Pi_m \boldsymbol{\theta} \in H_m$,

$$\langle \boldsymbol{f}, P\boldsymbol{f} \rangle_H = \boldsymbol{\theta}^\top \boldsymbol{A}_m \boldsymbol{\theta}.$$

The penalized estimator is defined as

$$\widehat{\boldsymbol{f}}_{T,\lambda} = \arg\min_{\boldsymbol{f} \in H_m} \left\{ S_{T,m}(\boldsymbol{f}) + \lambda_T \langle \boldsymbol{f}, P\boldsymbol{f} \rangle_H \right\},$$

where $\lambda_T > 0$ is a smoothing parameter. In practice, this is equivalent to solving

$$\widehat{\boldsymbol{f}}_{T,\lambda} = \Pi_m \widehat{\boldsymbol{\theta}}_{T,\lambda}, \quad \text{with} \quad \widehat{\boldsymbol{\theta}}_{T,\lambda} = \arg\min_{\boldsymbol{\theta} \in \mathbb{R}^{mJ}} \left\{ S_{T,m}(\Pi_m \boldsymbol{\theta}) + \lambda_T \boldsymbol{\theta}^\top \boldsymbol{A}_m \boldsymbol{\theta} \right\}.$$

The large-sample behavior of the penalized estimator is governed by the *effective degrees of freedom*, $d_{\text{eff}}(\lambda_T) := \operatorname{tr}(\boldsymbol{H}_\lambda)$, where the smoother matrix $\boldsymbol{H}_\lambda$ is given by

$$\boldsymbol{H}_\lambda = \boldsymbol{X}^\star \left(\boldsymbol{X}^{\star\top} \boldsymbol{X}^\star + \lambda_T T \boldsymbol{A}_m \right)^{-1} \boldsymbol{X}^{\star\top}.$$

Here, $\boldsymbol{X}^\star$ is the $T \times mJ$ design matrix of stacked high-frequency regressors, $\boldsymbol{A}_m$ is the $mJ \times mJ$ penalty matrix, and T is the low-frequency sample size.

Proposition 1. *1. If $\lambda_T \to 0$ and the effective degrees of freedom satisfy*

$$d_{\text{eff}}(\lambda_T)\,\log T = o(T),$$

then

$$\|\widehat{\boldsymbol{f}}_{T,\lambda} - \boldsymbol{\beta}\|_H \xrightarrow{p} 0 \qquad \text{as } T \to \infty.$$

2. If, in addition, $\sqrt{T}\lambda_T \to 0$, then for any linear functional $L(\boldsymbol{f}) = \langle \boldsymbol{f}, \boldsymbol{\psi}\rangle_H$ with fixed $\boldsymbol{\psi} \in H$,

$$\sqrt{T}\big(L(\widehat{\boldsymbol{f}}_{T,\lambda}) - L(\boldsymbol{\beta})\big) \xrightarrow{d} \mathcal{N}\left(0,\ \langle \boldsymbol{\psi}, Q^{-1}\Omega Q^{-1}\boldsymbol{\psi}\rangle_H\right),$$

where Q and Ω are the covariance and long-run variance operators defined in Theorem 1.

In applications, the condition

$$d_{\text{eff}}(\lambda_T)\,\log T = o(T)$$

is interpreted operationally by monitoring the effective degrees of freedom induced by the chosen penalty matrix $\boldsymbol{A}_m$ and smoothing parameter λ_T. Thus, the relevant notion of complexity is not the nominal number of within-period bins m, but the trace of the smoother matrix, $d_{\text{eff}}(\lambda_T) = \operatorname{tr}(\boldsymbol{H}_\lambda)$.

For standard second-difference spline penalties, $d_{\text{eff}}(\lambda_T)$ decreases monotonically with λ_T, so larger values of the smoothing parameter impose stronger regularization and reduce the effective dimension of the problem. In this sense, the asymptotic requirement in Proposition 1 provides practical guidance: even when m is large, stable estimation can be achieved by choosing λ_T so that the effective degrees of freedom remain small relative to $T/\log T$.

A proof of this proposition is provided in Appendix B. These results continue to hold even when the number of high-frequency bins m grows arbitrarily fast, provided the smoothing parameter is chosen so that the effective degrees of freedom increase sufficiently slowly relative to $T/\log T$. This shifts attention from the nominal dimension m to an operational notion of model complexity that can be monitored directly through $\operatorname{tr}(\boldsymbol{H}_\lambda)$. This observation naturally motivates a Bayesian formulation of the regularized U-MIDAS model, in which the smoothing penalty is interpreted as a Gaussian prior and uncertainty about economically relevant functionals can be quantified through posterior distributions.

3 Bayesian Inference for Regularized U-MIDAS

This section develops a Bayesian formulation of the regularized U-MIDAS model. We first show how the smoothing-spline penalty admits a natural interpretation as a Gaussian prior on the lag-weight coefficients and establish a Bernstein–von Mises result for linear functionals of the lag-weight function. We then describe a practical MCMC sampler for posterior computation.

Corollary 1. *Let $\boldsymbol{f} = \Pi_m \boldsymbol{\theta} \in H_m$ denote the unknown lag-weight function under the sieve representation, and let $L(\boldsymbol{f}) = \langle \boldsymbol{f}, \boldsymbol{\psi} \rangle_H$ for fixed $\boldsymbol{\psi} \in H$. Consider a Gaussian prior on $\boldsymbol{\theta}$ with precision matrix $\lambda_T \boldsymbol{A}_m$ (proper on* $\mathrm{Range}(\boldsymbol{A}_m)$ *and flat on* $\mathrm{Null}(\boldsymbol{A}_m)$*), and an independent inverse-Gamma prior on the noise variance, $\sigma^2 \sim \mathcal{IG}(a_0, b_0)$. Assume the conditions of Theorem 1 and $\sqrt{T}\lambda_T \to 0$. Then, as $T \to \infty$,*

$$\mathcal{L}\left(\sqrt{T}\big(L(\boldsymbol{f}) - L(\widehat{\boldsymbol{f}}_{T,\lambda})\big) \,\middle|\, \boldsymbol{y}\right) \xrightarrow{d} \mathcal{N}\left(0, \langle \boldsymbol{\psi}, Q^{-1}\boldsymbol{\psi} \rangle_H\right).$$

Combined with Proposition 1, this implies that

$$\mathcal{L}\left(\sqrt{T}\big(L(\boldsymbol{f}) - L(\boldsymbol{\beta})\big) \,\middle|\, \boldsymbol{y}\right) \xrightarrow{d} \mathcal{N}\left(0, \langle \boldsymbol{\psi}, Q^{-1}\Omega Q^{-1}\boldsymbol{\psi} \rangle_H\right).$$

The proof of this corollary follows standard arguments in the Bernstein–von Mises literature; the details are provided in Appendix C. The key step is a decomposition of the posterior distribution into a centered Gaussian fluctuation and the sampling distribution of the penalized estimator, together with the fact that the posterior variance of the fluctuation is asymptotically deterministic. Classical Gaussian shift arguments then imply asymptotic independence between the posterior center and spread (Castillo and Nickl, 2014).

While the argument is routine, the result is important for interpretation: Bayesian credible intervals for linear functionals $L(\boldsymbol{\beta})$, obtained from MCMC draws, are asymptotically equivalent to the frequentist confidence intervals derived in Proposition 1. This establishes the frequentist validity of posterior uncertainty quantification and justifies the use of Bayesian inference for economically meaningful summaries of the lag-weight function in applications.

Next, to describe the Bayesian MCMC sampler for the U-MIDAS model with cubic smoothing-spline regularization, we work with a discrete-time representation of the model. Let $\boldsymbol{y} = (y_1, \ldots, y_T)^\top$ denote the low-frequency response variable. For each $j = 1, \ldots, J$, let $x_{j,t}^{(m_j)}$ denote the jth high-frequency predictor, observed m_j times between $t-1$ and t. We allow the number of high-frequency bins m_j to vary across predictors.

Define the lag operator L^{c/m_j} by $L^{c/m_j} x_{j,t}^{(m_j)} = x_{j,t-c/m_j}^{(m_j)}$ for $c = 0, \ldots, m_j - 1$. The unrestricted MIDAS model can then be written as

$$y_t = \alpha + \sum_{j=1}^{J} \boldsymbol{x}_{j,t}^{\top} \boldsymbol{\beta}_j + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, \sigma^2),$$

where

$$\boldsymbol{x}_{j,t} = \left(x_{j,t}^{(m_j)}, x_{j,t-1/m_j}^{(m_j)}, \ldots, x_{j,t-(m_j-1)/m_j}^{(m_j)} \right)^{\top}.$$

This specification corresponds to an unrestricted MIDAS model with no parametric weighting function imposed on the high-frequency lags.

To incorporate smoothness across high-frequency bins, we model each coefficient vector $\boldsymbol{\beta}_j$ as the discretization of a smooth lag-weight function $\beta_j(\cdot)$ defined over continuous time $s \in [0, \tau]$. Specifically, we write

$$\beta_j(s) = (\Pi_{m_j} \boldsymbol{\beta}_j)(s),$$

and denote $\boldsymbol{\beta}_j = (\beta_j(0), \beta_j(\tau/m_j), \ldots, \beta_j((m_j - 1)\tau/m_j))^{\top}$.

We incorporate a standard cubic smoothing-spline regularization by penalizing the roughness of each lag-weight function. Specifically, the penalty takes the form

$$\lambda_j \int \left[\beta_j''(s) \right]^2 ds,$$

where $\lambda_j > 0$ is a smoothing parameter. In discrete form, this penalty is approximated by $\lambda_j \boldsymbol{\beta}_j^{\top} \boldsymbol{A}_{m_j} \boldsymbol{\beta}_j$, where $\boldsymbol{A}_{m_j}$ is the usual second-difference penalty matrix given by

$$\boldsymbol{A}_{m_j} = \boldsymbol{P}_{m_j}^{\top} \boldsymbol{P}_{m_j}, \qquad \boldsymbol{P}_{m_j} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{bmatrix}.$$

Under the Bayesian framework, we impose a Gaussian prior on each coefficient vector,

$$(\boldsymbol{\beta}_j \mid \tau_j^2) \sim \mathcal{N}\Big(\mathbf{0},\, \tau_j^2 \boldsymbol{A}_{m_j}^{-1} \Big),$$

which induces a smoothing-spline penalty in the posterior mode. This specification corresponds to a Gaussian smoothness prior of the type studied by Koop and Poirier (2004) and

Koop et al. (2005), where roughness is controlled through a second-difference precision structure. τ_j^2 is the spline variance parameter, which controls the overall shrinkage strength on $\boldsymbol{\beta}_j$. Here, $\tau_j^2 > 0$ is a predictor-specific spline variance parameter and should not be confused with the within-period horizon τ.

Conditional on the regression coefficients, the likelihood is given by

$$(\boldsymbol{y} \mid \alpha, \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_J, \sigma^2) \sim \mathcal{N}(\boldsymbol{m},\, \sigma^2 \boldsymbol{I}_T),$$

where $\boldsymbol{m} = (m_1, \ldots, m_T)^\top$ with $m_t = \alpha + \sum_{j=1}^{J} \boldsymbol{x}_{j,t}^\top \boldsymbol{\beta}_j$. As a result, the posterior mode for $\boldsymbol{\beta}_j$ coincides with the penalized least-squares estimator, with smoothing parameter $\lambda_j = \sigma^2/(T\tau_j^2)$. We complete the prior specification by setting $\alpha \sim \mathcal{N}(\mu_\alpha, V_\alpha)$, $\sigma^2 \sim \mathcal{IG}(a_0, b_0)$, and assigning independent inverse-Gamma priors to the spline variance parameters, $\tau_j^2 \sim \mathcal{IG}(a_\tau, b_\tau), j = 1, \ldots, J$.

Under these priors, the full conditional distribution of $\boldsymbol{\beta}_j$ is Gaussian,

$$\boldsymbol{\beta}_j \mid \cdot \sim \mathcal{N}(\widehat{\boldsymbol{\beta}}_j, \boldsymbol{D}_{\beta,j}),$$

where

$$\boldsymbol{D}_{\beta,j} = \left(\sum_{t=1}^{T} \boldsymbol{x}_{j,t} \boldsymbol{x}_{j,t}^\top / \sigma^2 + \boldsymbol{A}_{m_j} / \tau_j^2 \right)^{-1}, \quad \widehat{\boldsymbol{\beta}}_j = \boldsymbol{D}_{\beta,j} \left(\sum_{t=1}^{T} \boldsymbol{x}_{j,t} \Big(y_t - \alpha - \sum_{l \neq j} \boldsymbol{x}_{l,t}^\top \boldsymbol{\beta}_l \Big) / \sigma^2 \right).$$

Here and throughout, the notation "$\mid \cdot$" denotes conditioning on all other parameters and the observed data. The full conditional distribution of the intercept α is

$$\alpha \mid \cdot \sim \mathcal{N}(\hat{\alpha}, D_\alpha),$$

where

$$D_\alpha = (T/\sigma^2 + 1/V_\alpha)^{-1}, \qquad \hat{\alpha} = D_\alpha \left(\sum_{t=1}^{T} \big(y_t - \sum_{j=1}^{J} \boldsymbol{x}_{j,t}^\top \boldsymbol{\beta}_j \big) / \sigma^2 + \mu_\alpha / V_\alpha \right).$$

Finally, the conditional posterior distributions of σ^2 and τ_j^2 are given by

$$\sigma^2 \mid \cdot \sim \mathcal{IG}\left(\tfrac{T}{2} + a_0,\ \tfrac{1}{2} \sum_{t=1}^{T} \epsilon_t^2 + b_0 \right), \qquad \tau_j^2 \mid \cdot \sim \mathcal{IG}\big(\tfrac{m_j}{2} + a_\tau,\ \tfrac{1}{2} \boldsymbol{\beta}_j^\top \boldsymbol{A}_{m_j} \boldsymbol{\beta}_j + b_\tau \big).$$

4 Monte Carlo Experiments

This section reports Monte Carlo experiments designed to illustrate the asymptotic results developed in Sections 2 and 3. The experiments serve two purposes. First, they verify that unregularized U-MIDAS performs well when the number of high-frequency bins grows moderately relative to the low-frequency sample size. Second, they show that under extreme frequency mismatch, unrestricted estimation becomes unstable, while smoothing-spline regularization restores stability and supports inference on linear functionals.

4.1 Simulation Design and Evaluation

Data-generating process. We simulate data from the continuous-time model introduced in Section 2. Let $\boldsymbol{X}(s) \in \mathbb{R}^J$ satisfy

$$d\boldsymbol{X}(s) = \boldsymbol{\mu}\, ds + \boldsymbol{\Sigma}\, d\boldsymbol{W}(s),$$

where $\boldsymbol{W}(s)$ is a J-dimensional standard Brownian motion. The low-frequency outcome is generated as

$$y_t = \alpha + \int_0^\tau \boldsymbol{\beta}(s)^\top d\boldsymbol{X}(t-s) + u_t, \qquad t = 1, \dots, T,$$

where $\boldsymbol{\beta}(\cdot) \in L^2([0,\tau]; \mathbb{R}^J)$ and $\mathbb{E}[u_t \mid \mathcal{F}_{t^-}^X] = 0$. To obtain an estimable discrete regression, we partition $[0,\tau]$ into m equal bins of width $h = \tau/m$ and define the increments

$$\Delta \boldsymbol{X}_{t,\ell} = \boldsymbol{X}(t-(\ell-1)h) - \boldsymbol{X}(t-\ell h), \qquad \ell = 1, \dots, m,$$

so that the binned U-MIDAS regression takes the form

$$y_t = \alpha + \sum_{\ell=1}^{m} \boldsymbol{\theta}_\ell^\top \Delta \boldsymbol{X}_{t,\ell} + u_t, \qquad \boldsymbol{\theta}_\ell = \boldsymbol{\beta}((\ell-1)h).$$

Parameter settings. We set $\tau = 1$ and consider $T \in \{100, 200\}$ and $J \in \{1, 3\}$. Throughout, we set $\boldsymbol{\mu} = \mathbf{0}$. For $J = 1$ we take $\boldsymbol{\Sigma} = 1$. For $J = 3$ we set

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 & 0 \\ 0.3 & 1 & 0 \\ 0.2 & 0.4 & 1 \end{pmatrix}.$$

We consider two specifications for the true lag-weight function:

$$\begin{aligned}\text{(Smooth)} \quad & \boldsymbol{\beta}(s) = c\exp(-6s)\big(1 + 0.5\sin(4\pi s)\big),\\ \text{(Localized)} \quad & \boldsymbol{\beta}(s) = c\exp\left(-\frac{(s-0.25)^2}{2(0.05)^2}\right),\end{aligned}$$

with the scale c chosen to yield a moderate signal-to-noise ratio. The low-frequency disturbance follows

$$u_t = \rho u_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \qquad \rho \in \{0, 0.5\}.$$

Frequency mismatch regimes. To connect directly to the theory, we consider both moderate and extreme growth of m:

$$m \in \Big\{\lfloor 2\sqrt{T}\rfloor,\ \lfloor 5\sqrt{T}\rfloor,\ \lfloor T^{0.75}\rfloor,\ \lfloor 2T\rfloor\Big\}.$$

The first three choices satisfy $m = o(T/\log T)$, while $m = 2T$ represents an extreme mismatch regime in which unrestricted estimation is expected to be unstable.

Estimators. We compare the following estimators: (i) unregularized U-MIDAS (OLS); (ii) adaptive group LASSO shrinkage, which introduces L1-norm penalty on each $\boldsymbol{\beta}_j$ as a group (Wang and Leng, 2008); (iii) cubic smoothing-spline regularization with a fixed smoothing parameter $\lambda_T = T^{-3/4}$ (chosen to satisfy the conditions in Proposition 1); and (iv) cubic smoothing-spline regularization with an adaptive smoothing parameter, implemented through the hierarchical prior on τ_j^2 so that $\lambda_j = \sigma^2/(T\tau_j^2)$ varies across MCMC iterations.

Performance measures. We summarize estimation accuracy using the integrated squared error (ISE),

$$\mathrm{ISE}(\widehat{\boldsymbol{f}}) = \int_0^\tau \|\widehat{\boldsymbol{f}}(s) - \boldsymbol{\beta}(s)\|^2\, ds \approx \sum_{\ell=1}^{m} h\, \|\widehat{\boldsymbol{\theta}}_\ell - \boldsymbol{\theta}_\ell^0\|^2,$$

where $\widehat{\boldsymbol{f}}(s)$ is the step-function estimate implied by the binned coefficients and $\boldsymbol{\theta}_\ell^0$ denotes the true bin coefficient. We also report numerical stability measures, including the condition number of $\boldsymbol{X}^{\star\top}\boldsymbol{X}^\star$ and the Monte Carlo variability of the estimated coefficients across replications.

Implementation. Each design is replicated $R = 500$ times. For each replication, increments are simulated as

$$\Delta \boldsymbol{X}_{t,\ell} = \boldsymbol{\mu} h + \boldsymbol{\Sigma} \Delta \boldsymbol{W}_{t,\ell}, \qquad \Delta \boldsymbol{W}_{t,\ell} \sim \mathcal{N}(\boldsymbol{0}, h \boldsymbol{I}_J),$$

and the outcome $\{y_t\}$ is constructed accordingly. For Bayesian estimators, we run the MCMC sampler described in Section 3 and summarize results using posterior means and credible intervals.

4.2 Simulation Results

We report here results for Case 2 (multivariate, $J = 3$) with $T = 200$, which provides the clearest illustration of the effects of extreme frequency mismatch. Additional results for the univariate case ($J = 1$) and for smaller sample sizes are reported in Appendix D.

Across all designs, we consider four estimators: (i) unregularized U-MIDAS (OLS), (ii) adaptive group LASSO, (iii) cubic smoothing-spline regularization with a fixed smoothing parameter $\lambda_T = T^{-3/4}$ (chosen to satisfy the conditions in Proposition 1), and (iv) cubic smoothing-spline regularization with an adaptive smoothing parameter implemented through the hierarchical prior on τ_j^2, so that $\lambda_j = \sigma^2/(T\tau_j^2)$ varies across MCMC iterations. For each design we report the integrated squared error (ISE) of the implied step-function estimator, the average Monte Carlo variance of the estimated coefficients across bins, and the mean condition number of $\boldsymbol{X}^{\star\top}\boldsymbol{X}^{\star}$ across replications.

Table 1 contrasts unregularized and regularized U-MIDAS as the number of within-period bins m increases. Under moderate growth of m (the first three columns), all estimators perform comparably in terms of ISE. Under extreme mismatch ($m = 2T$), however, the unregularized U-MIDAS estimator becomes numerically unstable: the ISE and the variance of the estimated coefficients increase dramatically, and the condition number of $\boldsymbol{X}^{\star\top}\boldsymbol{X}^{\star}$ explodes. Adaptive group LASSO mitigates some of this instability, but its performance still deteriorates when m is large relative to T.

In contrast, cubic smoothing-spline regularization remains stable even under extreme mismatch. Both fixed and adaptive choices of the smoothing parameter yield substantially smaller ISE and coefficient variance, with the adaptive specification providing modest additional gains in the most severe cases. These findings are consistent with Proposition 1, which shows that smoothing-spline regularization delivers consistent estimation and valid inference for linear functionals as long as the effective degrees of freedom grow sufficiently

slowly, regardless of how fast m increases.

Next, we also investigate the empirical coverage rate of 95% highest posterior density credible intervals of linear functionals under cubic smoothing spline regularization and different degrees of frequency mismatch. Relevant results are illustrated in Appendix D. Simulation results show that cubic smoothing spline leads to stable coverage of credible intervals even under extreme frequency mismatch.

Table 1. Results of Monte Carlo Experiments under Case 2 (Multivariate) with $T = 200$

$m =$	$2\sqrt{T}$		$5\sqrt{T}$		$T^{0.75}$		$2T$	
$\rho =$	0	0.5	0	0.5	0	0.5	0	0.5
True lag-weight function: Smooth								
No shrinkage								
ISE	0.800	1.051	720.960	724.829	4.429	5.814	912.569	912.630
Mean variance of $\hat{\boldsymbol{\beta}}$	0.267	0.351	240.315	241.607	1.477	1.939	302.998	303.018
Adaptive group LASSO								
ISE	0.633	0.786	2.044	2.255	1.431	1.650	5.152	5.212
Mean variance of $\hat{\boldsymbol{\beta}}$	0.205	0.252	0.597	0.646	0.439	0.498	0.463	0.481
Cubic smoothing spline (fixed λ_T)								
ISE	0.168	0.217	0.344	0.449	0.266	0.348	1.704	2.225
Mean variance of $\hat{\boldsymbol{\beta}}$	0.056	0.072	0.115	0.150	0.089	0.116	0.568	0.742
Cubic smoothing spline								
ISE	0.097	0.117	0.116	0.142	0.101	0.122	0.334	0.412
Mean variance of $\hat{\boldsymbol{\beta}}$	0.028	0.034	0.038	0.046	0.032	0.038	0.111	0.137
True lag-weight function: Localized								
No shrinkage								
ISE	0.800	1.051	720.871	724.799	4.429	5.814	911.041	911.102
Mean variance of $\hat{\boldsymbol{\beta}}$	0.267	0.351	240.283	241.594	1.477	1.939	302.877	302.897
Adaptive group LASSO								
ISE	0.562	0.685	1.522	1.670	1.130	1.295	3.433	3.490
Mean variance of $\hat{\boldsymbol{\beta}}$	0.178	0.215	0.431	0.461	0.336	0.376	0.319	0.336
Cubic smoothing spline (fixed λ_T)								
ISE	0.225	0.275	0.346	0.452	0.272	0.354	1.704	2.225
Mean variance of $\hat{\boldsymbol{\beta}}$	0.060	0.077	0.115	0.150	0.089	0.117	0.568	0.742
Cubic smoothing spline								
ISE	0.249	0.309	0.234	0.292	0.229	0.287	0.345	0.425
Mean variance of $\hat{\boldsymbol{\beta}}$	0.073	0.089	0.070	0.085	0.067	0.082	0.115	0.141
Mean $\kappa(\boldsymbol{X}^{\star\top}\boldsymbol{X}^{\star})$	249	249	1.4×10^{19}	1.4×10^{19}	4561	4561	1.2×10^{21}	1.2×10^{21}

5 Nowcasting US Real GDP

In this section, we apply the proposed regularized U-MIDAS framework to nowcasting US quarterly real GDP growth. The application builds on the large-scale MIDAS setup in

Section 5 of Chan et al. (2025), augmented with additional aggregate corporate variables—including a recession indicator (monthly), total debt (daily), and net income (daily)—from Pettenuzzo et al. (2024). The full set of predictors and their transformations are summarized in Table 2.

Table 2. List of Variables Used in the Nowcasting Application

Variable	Frequency	Transformation
Real GDP	Quarterly	$400\Delta \ln x_t$
NFCI	Weekly	Level
Interest Rate Spread	Daily	Level
Industrial Production	Monthly	$100\Delta \ln x_t$
Housing Starts	Monthly	$100\Delta \ln x_t$
Average Weekly Hours of Production	Monthly	$\ln x_t$
Civilian Labor Force Level	Monthly	$100\Delta \ln x_t$
All Employees, Total Nonfarm	Monthly	$100\Delta \ln x_t$
Capacity Utilization	Monthly	$100\Delta \ln x_t$
Unemployment Rate	Monthly	Level
CPI Inflation	Monthly	$100\Delta \ln x_t$
S&P 500	Daily	$100\Delta \ln x_t$
Fed Funds Effective Rate	Daily	Level
BAA Corporate Bond Yield	Daily	Level
US/UK Exchange Rate	Daily	$100\Delta \ln x_t$
VIX	Daily	$\ln x_t$
Recession Indicator	Monthly	Level
Net income	Daily	Level
Total debt	Daily	Level

We adopt a recursive nowcasting scheme. Because daily VIX data are available only from 1990 onward, the initial in-sample period runs from 1991Q1 to 2004Q4, with the out-of-sample evaluation period covering 2005Q1–2024Q4. Forecast accuracy is assessed using the root mean squared forecast error (RMSFE) and the success ratio, which measures directional accuracy—the fraction of periods in which the nowcast correctly predicts the sign of GDP growth changes (MacCarthy and Snudden, 2024).

We compare five specifications. The first estimates a standard U-MIDAS model without regularization. The second and third use naive temporal aggregation, averaging daily predictors to weekly and monthly frequency, respectively. The final two specifications estimate U-MIDAS models with adaptive group LASSO shrinkage and with cubic smoothing-spline regularization.

Table 3 reports the nowcasting results. The comparison across the first three specifications highlights the effect of frequency mismatch: as the degree of mismatch increases, the perfor-

mance of unregularized U-MIDAS deteriorates sharply, with very large RMSFEs and poor directional accuracy. Naively aggregating high-frequency data alleviates some instability, but at the cost of discarding potentially informative within-period variation.

Introducing regularization substantially improves performance. Both adaptive group LASSO and cubic smoothing-spline regularization yield much lower RMSFEs and higher success ratios than the unregularized and aggregated benchmarks. Overall, cubic smoothing-spline regularization delivers forecast accuracy comparable to adaptive group LASSO while maintaining a smoother and more interpretable lag structure.

Table 3. Nowcast Results under the U-MIDAS Model

	2005Q1-2024Q4		2005Q1-2019Q4	
	RMSFE	Success Ratio	RMSFE	Success Ratio
No shrinkage	308.6803	0.4875	222.8082	0.4667
No shrinkage (weekly averages)	33.2979	0.5750	37.6428	0.6000
No shrinkage (monthly averages)	6.2577	0.6875	3.6310	0.6833
Adaptive group LASSO	4.6051	0.7125	2.5086	0.7000
Cubic smoothing spline	5.1323	0.7125	2.5245	0.7333

Adaptive group LASSO performs well in this application despite lacking theoretical justification under extreme frequency mismatch, reflecting its focus on predictor-level shrinkage rather than smooth lag recovery. As shown in the Monte Carlo experiments, however, its performance can deteriorate when m is very large relative to T, whereas smoothing-spline regularization remains stable as long as the effective degrees of freedom are controlled in accordance with Proposition 1.

Figure 1 illustrates differences in the implied lag structure for selected weekly and daily predictors in the 2019Q4 nowcast (in-sample period 1991Q1–2019Q3). Under naive aggregation, coefficients are constant within weekly or monthly blocks, while both shrinkage methods produce smoother profiles across high-frequency bins.

A direct comparison in Figure 2 shows that cubic smoothing-spline regularization yields smooth, gradually decaying lag effects—particularly for NFCI, the federal funds rate, net income, and total debt—whereas adaptive group LASSO produces more irregular patterns.

To connect to the theoretical emphasis on inference for linear functionals, we additionally examine the empirical distribution of $\sqrt{T}\big(L(\widehat{\boldsymbol{f}}) - L(\boldsymbol{\beta})\big)$, where $L(\boldsymbol{f}) = \langle \boldsymbol{f}, \boldsymbol{\psi} \rangle_H$, for representative weighting functions $\boldsymbol{\psi}$ that emphasize early-lag, late-lag, and oscillatory patterns. Figure 3 reports posterior inference for standardized linear functionals under economically

meaningful weighting schemes.[1] Across specifications, posterior distributions are close to standard normal and 95% highest posterior density intervals exhibit stable coverage, illustrating the practical relevance of inference for linear functionals of the lag-weight function.

[1]We consider an early-lag weighting function $\psi(s) = \exp(-2s)$, which imposes higher weights on more recent observations, and a late-lag weighting scheme $\psi(s) = 1 - \exp(-2s)$, which puts higher weights on higher-order lags. Also, we consider an oscillatory weighting function $\psi(s) = \cos(2\pi s)$.

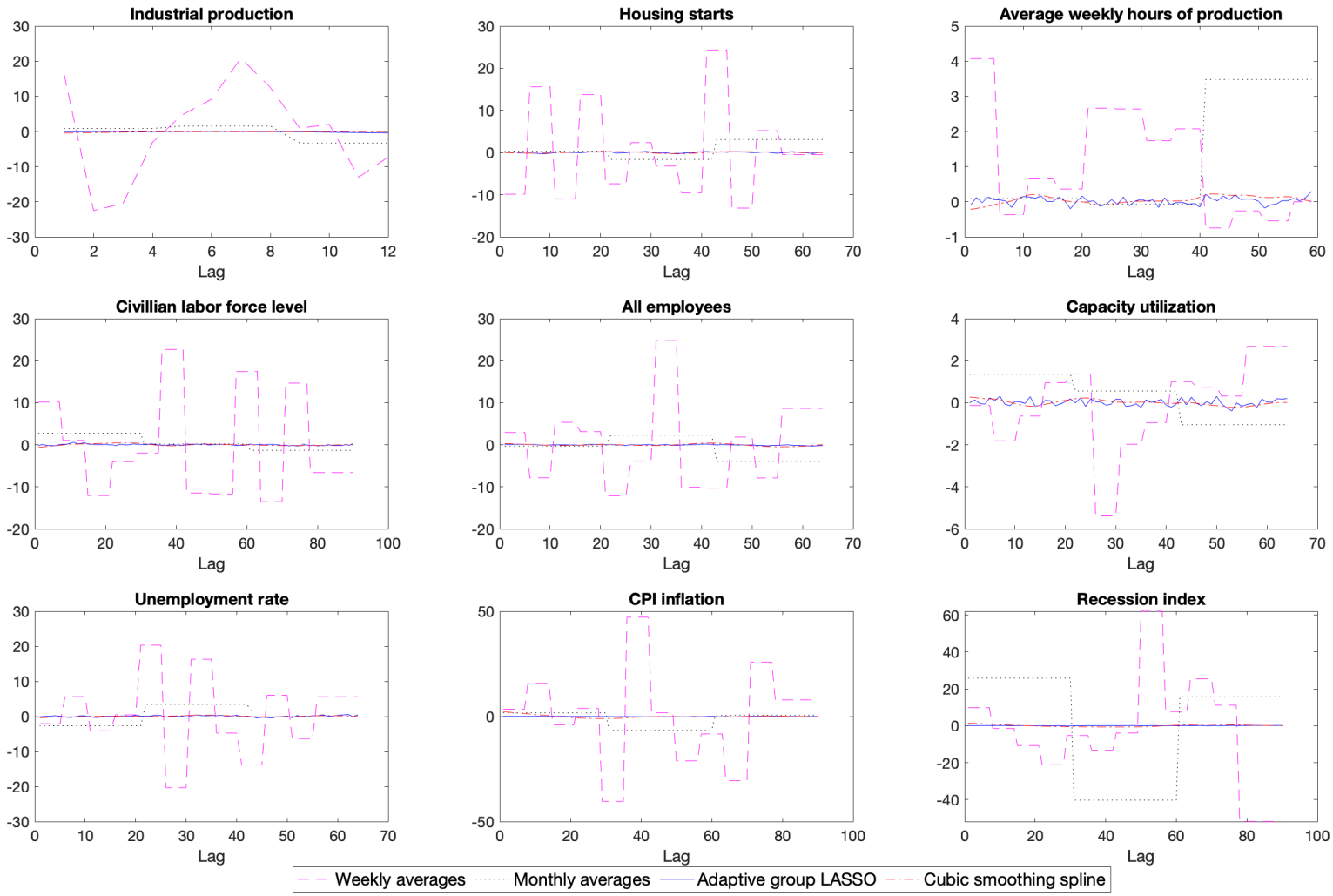

Figure 1. Estimated posterior means of coefficients based on the in-sample period 1991Q1-2019Q3

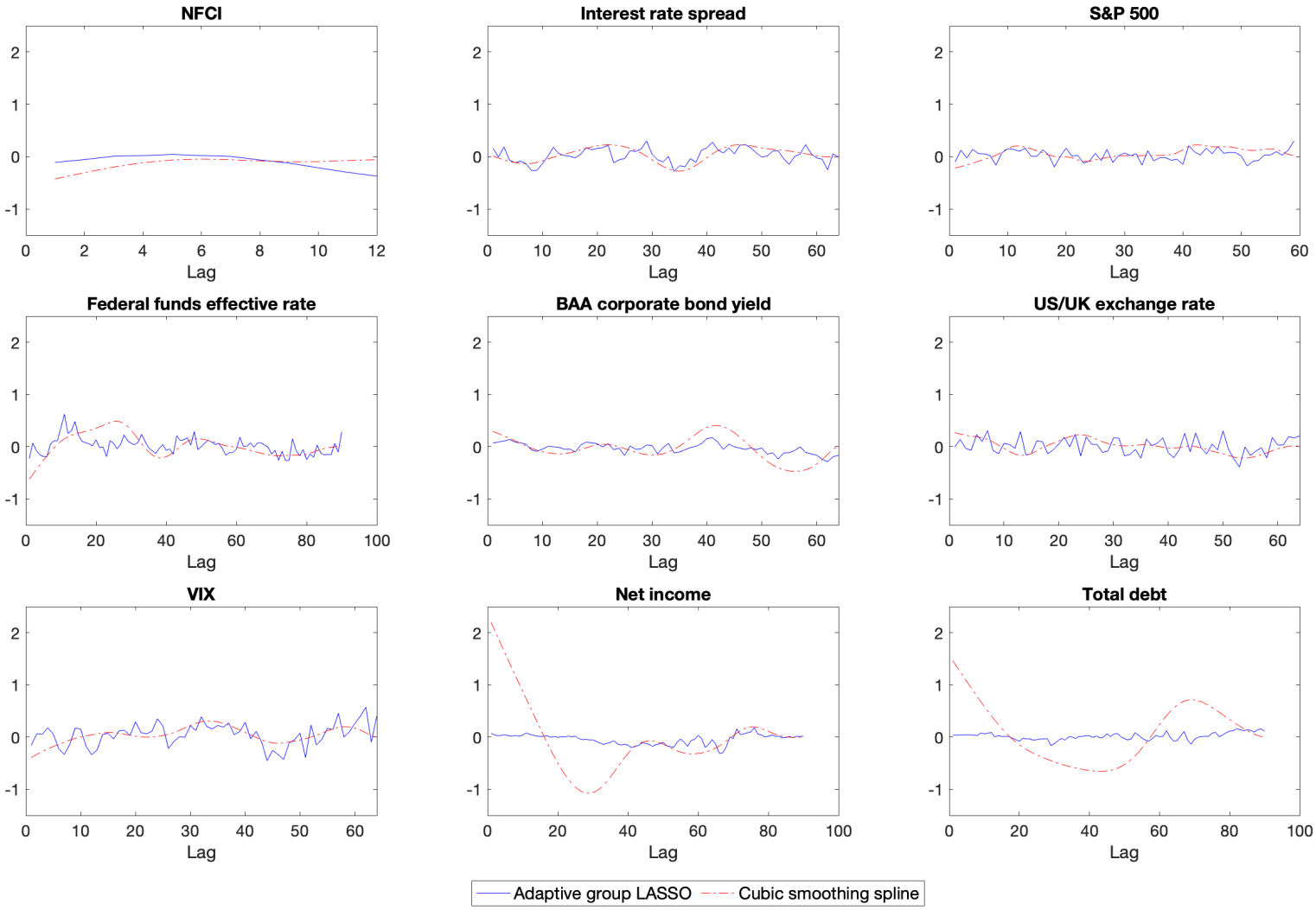

Figure 2. Estimated posterior means of coefficients based on the in-sample period 1991Q1-2019Q3 under shrinkage approaches

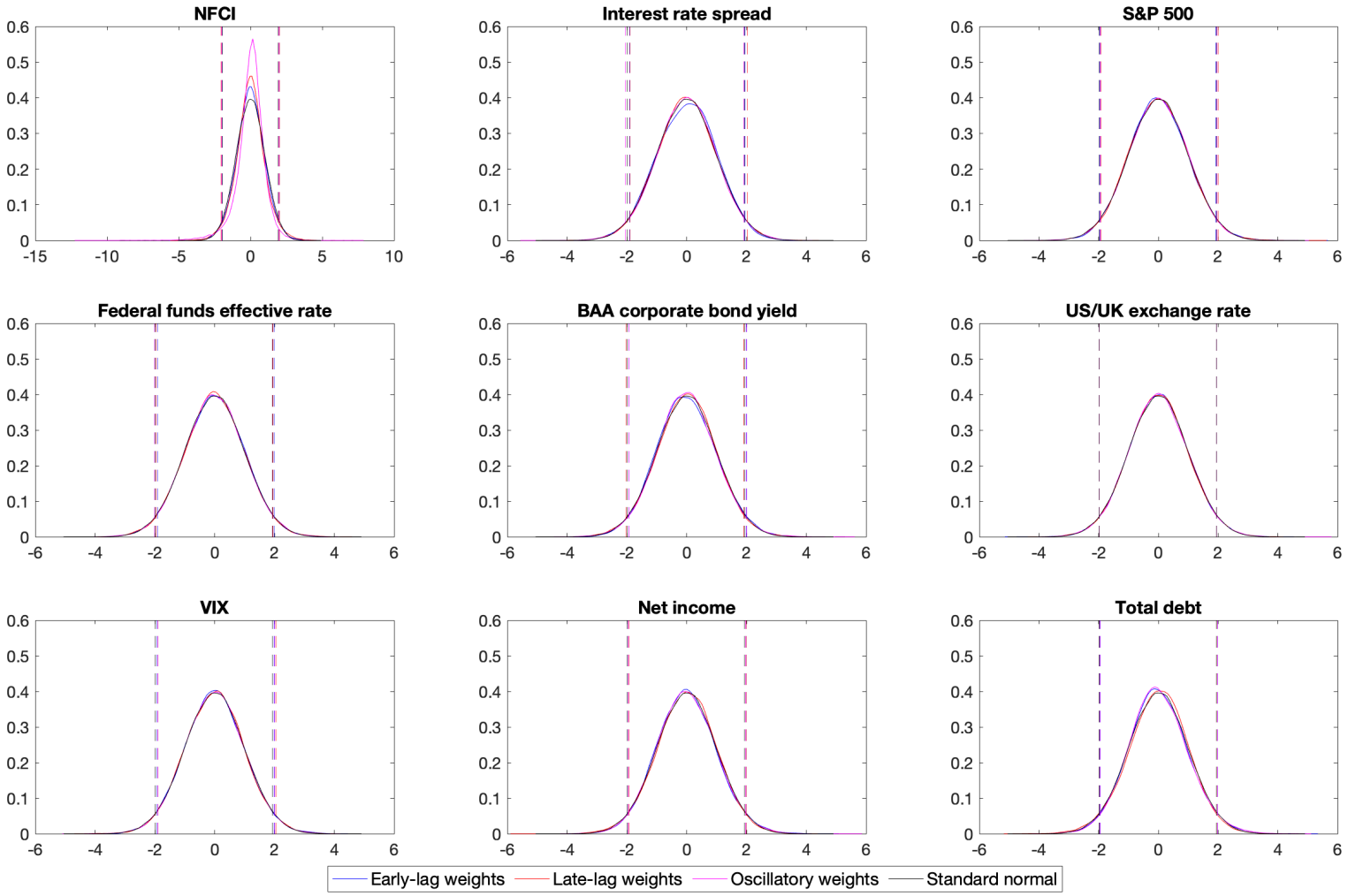

Figure 3. Posterior distributions of $L(\hat{f})$ based on the in-sample period 1991Q1-2019Q3

6 Measuring Macro Uncertainty from the Survey of Professional Forecasters

This section examines whether high-frequency financial information helps explain and forecast macroeconomic uncertainty, measured by disagreement in the Survey of Professional Forecasters (SPF).[2] Our sample spans 1991Q1–2019Q3. We consider SPF forecasts of (i) real GDP growth, (ii) CPI inflation, and (iii) the unemployment rate, and construct disagreement at the one-quarter-ahead forecast horizon ($h = 1$). Throughout, the dependent variable is the cross-sectional dispersion of SPF point forecasts for the corresponding target.

6.1 Construction of the SPF Disagreement Series

Let $t \in \{1, \ldots, T\}$ index quarters. For a given target variable, let $f_{i,t}^{(h)}$ denote the h-step-ahead SPF forecast submitted by forecaster $i \in \{1, \ldots, N_t\}$ in quarter t, where N_t is the number of respondents in that quarter. We define the SPF *forecast disagreement* series as the cross-sectional dispersion of forecasts,

$$y_t := \mathrm{Disp}_t \left(f_{1,t}^{(h)}, \ldots, f_{N_t,t}^{(h)} \right), \tag{1}$$

where $\mathrm{Disp}_t(\cdot)$ is a dispersion operator. In the empirical analysis, we take $\mathrm{Disp}_t(\cdot)$ to be the cross-sectional standard deviation,

$$y_t = \sqrt{\frac{1}{N_t - 1} \sum_{i=1}^{N_t} \left(f_{i,t}^{(h)} - \bar{f}_t^{(h)} \right)^2}, \qquad \bar{f}_t^{(h)} \equiv \frac{1}{N_t} \sum_{i=1}^{N_t} f_{i,t}^{(h)}. \tag{2}$$

This construction yields a continuous quarterly measure of forecast disagreement, with larger values of y_t indicating greater heterogeneity in professional forecasts. We interpret SPF disagreement as a proxy for macroeconomic uncertainty. While disagreement does not coincide with the representative-agent conditional variance, it captures dispersion in beliefs and information sets and is therefore a natural real-time measure of perceived uncertainty. This approach is widely used in the macro uncertainty and expectations literature as a timely and informative indicator of uncertainty.

[2]The SPF is a quarterly survey conducted by the Federal Reserve Bank of Philadelphia that collects forecasts from a panel of professional forecasters for a broad set of macroeconomic variables.

6.2 Intra-Quarter Timing and Financial Information

We relate the quarterly SPF disagreement series $\{y_t\}_{t=1}^{T}$ to high-frequency financial covariates observed within quarter t. Let $\boldsymbol{\beta}(\cdot) = (\beta_1(\cdot), \ldots, \beta_J(\cdot))^\top$ denote the vector of lag-weight functions assoicated with the high-frequency variables. Our objective is to assess how intra-quarter innovations in each high-frequency variable map into macroeconomic uncertainty as perceived by professional forecasters.

Under the U-MIDAS specification, high-frequency innovations in predictor j affect SPF disagreement through the predictor-specific lag-weight function $\beta_j(\cdot)$. Rather than focusing on pointwise estimates of $\beta_j(\cdot)$, which are high dimensional and difficult to interpret, we emphasize inference on low-dimensional, economically meaningful *linear functionals* of $\beta_j(\cdot)$. These functionals summarize the timing and overall importance of intra-quarter information in a transparent way.

Formally, for predictor j we consider linear functionals of the form

$$L_j(\beta_j) := \int_0^\tau \beta_j(s)\, \psi(s)\, ds,$$

where $\psi(\cdot)$ is a user-chosen weighting function. Because the model loads on high-frequency innovations of the financial process, these functionals directly characterize how shocks arriving at different points within the quarter contribute to SPF disagreement.

To summarize $\beta_j(\cdot)$ in economically meaningful ways, we focus on three linear functionals corresponding to different notions of timing.

1. Total (cumulative) effect. Setting $\psi_{\text{tot}}(s) = 1$, $s \in [0, \tau]$, we have

$$L_{j,\text{tot}} = \int_0^\tau \beta_j(s)\, ds,$$

which measures the overall contribution of predictor j aggregated over the entire quarter.

2. Late-quarter information content. To isolate information arriving near the end of the quarter, we fix a cutoff $c \in (0, \tau)$ and define $\psi_{\text{late}}(s) = \mathbf{1}\{s \in [0, c]\}$, yielding

$$L_{j,\text{late}} = \int_0^c \beta_j(s)\, ds.$$

In the empirical application, we set $c = \tau/3$, so that this functional captures the effect of the final third of intra-quarter observations.

3. Smooth end-of-quarter tilt. Finally, to obtain a smooth measure that places greater weight on more recent information, we consider $\psi_{\text{tilt}}(s) = 1 - s/\tau$, which implies

$$L_{j,\text{tilt}} = \int_0^\tau \beta_j(s) \left(1 - \frac{s}{\tau}\right) ds.$$

Relative to $L_{j,\text{tot}}$, this functional emphasizes late-quarter information while retaining a smooth weighting scheme.

6.3 Empirical Results

The empirical analysis addresses three questions. First, *when* within a quarter do financial variables contain the most information about subsequent macroeconomic uncertainty—early in the quarter, late in the quarter, or throughout? Second, *which* financial predictors exhibit the strongest systematic relationship with SPF disagreement once high-frequency information is aggregated through the estimated lag-weight functions? Third, *how precisely* can these economically meaningful objects be estimated, as reflected by posterior credible intervals for the associated linear functionals.

We report results for macroeconomic uncertainty measured by SPF disagreement in real GDP growth. Results for CPI inflation and the unemployment rate are qualitatively similar and are reported in Appendix E. The analysis uses quarterly data from 1991Q1–2019Q3 and the same set of high-frequency financial predictors listed in Table 2, which covers the variables most commonly used in the macro uncertainty literature (e.g., Rossi and Sekhposyan, 2015; Jurado et al., 2015). Posterior inference is based on 150,000 MCMC iterations, with the first 50,000 discarded as burn-in.

Figure 4 plots the posterior mean estimates of the predictor-specific lag-weight functions under cubic smoothing-spline regularization. For several key predictors—such as NFCI, BAA corporate bond yields, net income, and total debt—the estimated effects display a smooth decay as lag order increases, consistent with economically plausible intra-quarter timing patterns. While the decay is not strictly monotonic, the overall shape is stable and similar across uncertainty measures based on GDP growth and unemployment forecasts.

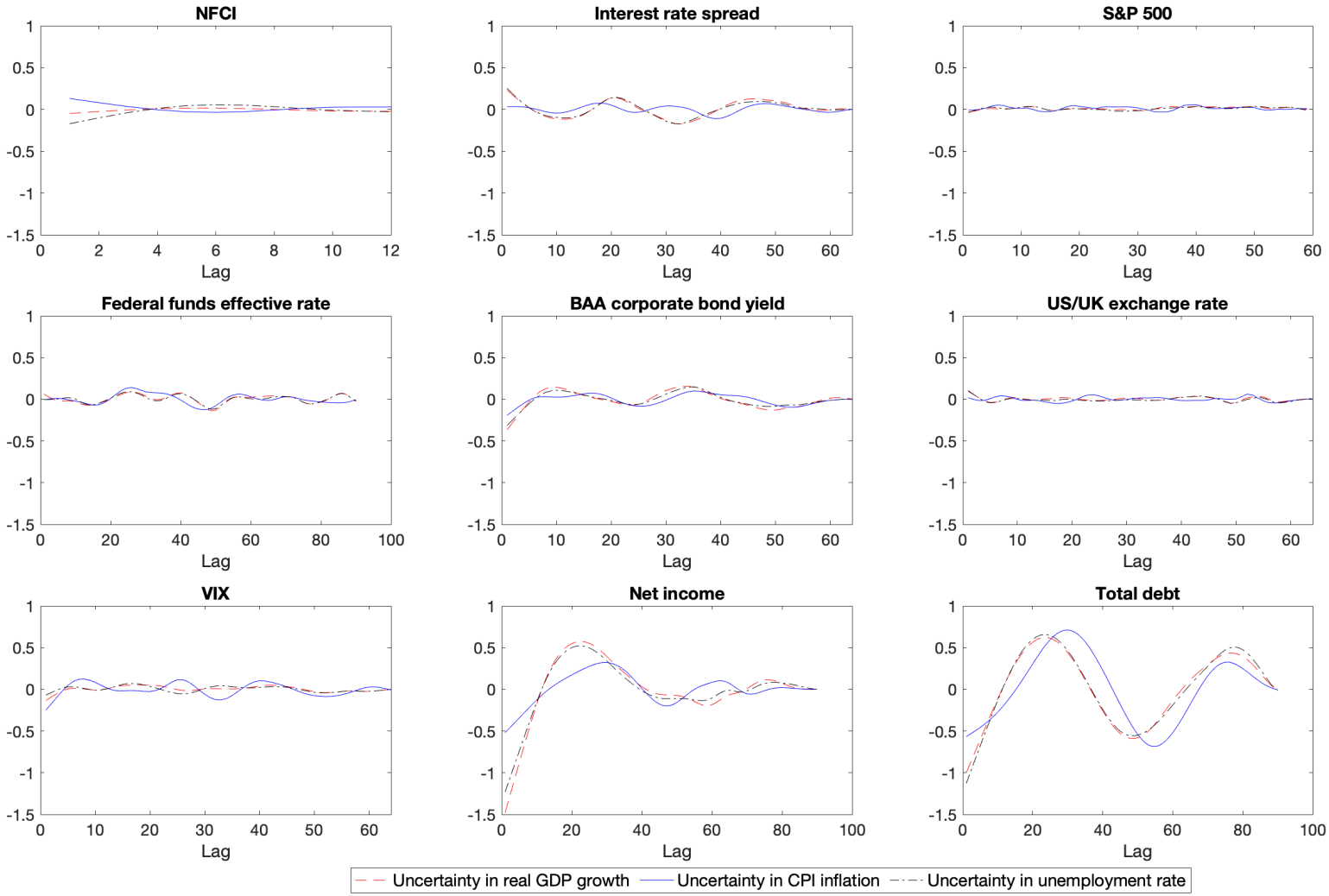

Figure 4. Estimated posterior means of coefficients, 1991Q1-2019Q3

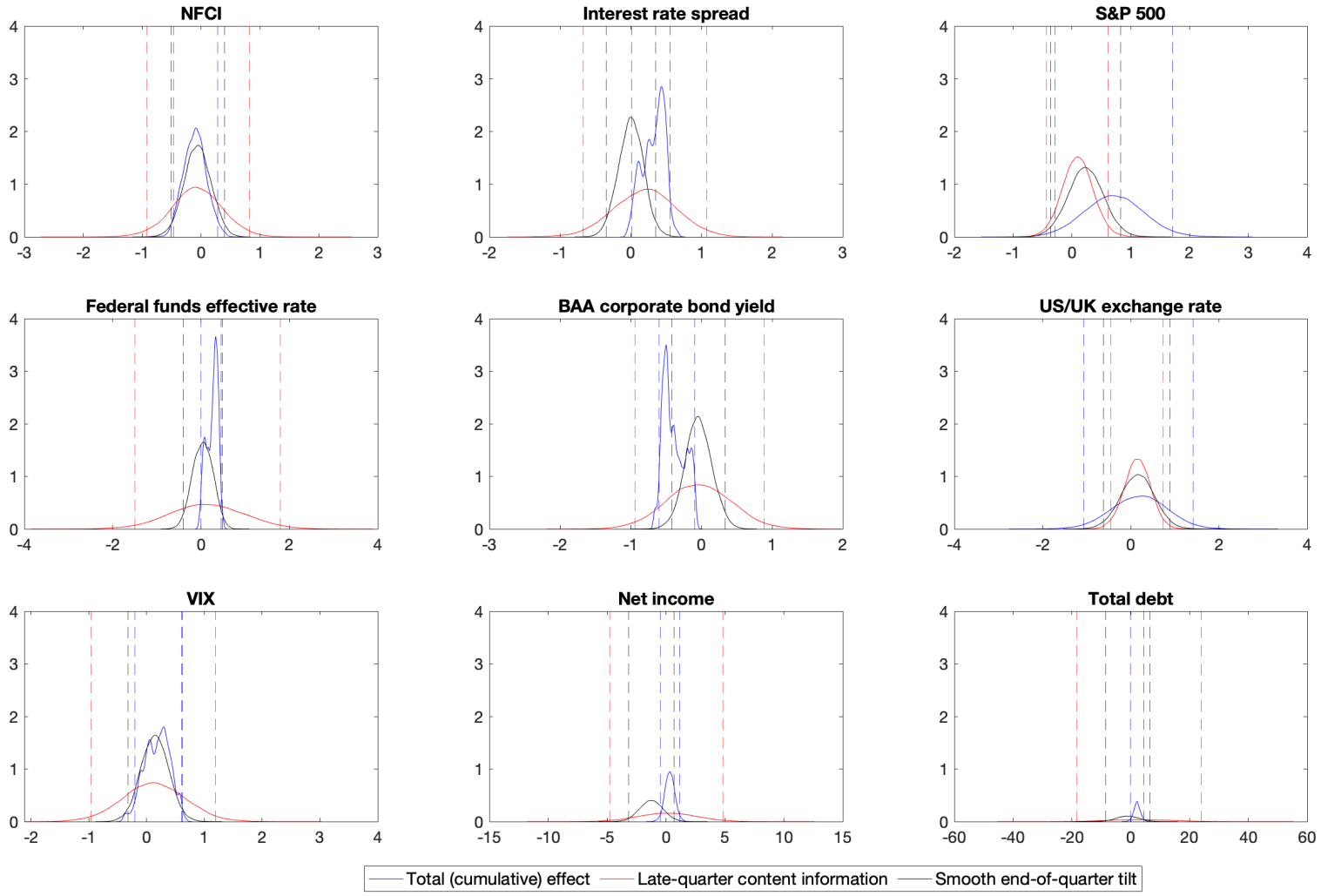

Figure 5. Uncertainty on real GDP growth: Posterior distributions of $L_{\text{tot}}(\beta)$, $L_{\text{late}}(\beta)$, and $L_{\text{tilt}}(\beta)$

Turning to inference on linear functionals, Figure 5 reports posterior distributions for three economically interpretable summaries: the total (cumulative) effect, the late-quarter effect, and a smooth end-of-quarter tilt. These objects correspond directly to the linear functionals covered by our asymptotic theory. Under the regularized U-MIDAS specification, posterior credible intervals for fixed weighting functions are asymptotically valid (Corollary 1), providing a principled basis for uncertainty quantification of within-quarter timing effects. In finite samples, posterior shapes may depart slightly from Gaussianity, so we report highest posterior density intervals in addition to posterior means in Appendix E.

Across predictors, late-quarter functionals are typically estimated less precisely than cumulative or smoothly weighted functionals, reflecting the smaller effective information set near the end of the quarter. At the same time, posterior means often indicate that late-quarter information has a larger marginal impact on uncertainty than earlier within-quarter contributions. This pattern suggests that early and late intra-quarter effects may partially offset each other, so that the total functional captures a net cumulative effect rather than the full magnitude of within-quarter variation.

Taken together, the evidence indicates that financial information arriving late in the quarter tends to have the strongest marginal association with macroeconomic uncertainty, particularly for credit-related variables, but is estimated with greater uncertainty. Cumulative and smoothly weighted summaries provide more precise inference and often capture economically meaningful net effects after early- and late-quarter contributions offset each other. These findings illustrate why inference on linear functionals—rather than pointwise lag coefficients—is both economically informative and statistically well behaved under extreme frequency mismatch.

7 Concluding Remarks and Future Research

This paper studies unrestricted MIDAS regressions when the frequency mismatch is very large. In that case, the instability of U-MIDAS can be viewed as an ill-posed problem of recovering the lag-weight function. Using a continuous-time representation, we show when unrestricted estimation is still valid and when regularization becomes necessary. We then use smoothing-spline regularization to control complexity through the effective degrees of freedom. This yields consistent estimation and valid inference for fixed linear functionals, even when the number of high-frequency observations is large. The Bayesian version is convenient in practice and produces credible intervals with good frequentist coverage.

The empirical results also support focusing on linear functionals rather than individual lag coefficients. In the GDP nowcasting exercise and in the SPF disagreement application, regularized U-MIDAS gives stable estimates and sensible summaries of within-quarter timing effects. In both cases, information arriving later in the quarter often matters more, while cumulative and smoothly weighted functionals are estimated more precisely. Overall, the framework is useful for extracting information from high-frequency data without over-interpreting noisy lag-by-lag estimates.

There are several directions for future work. One is to allow the lag-weight function to vary over time, or to let the amount of regularization depend on the state of the economy. Another is to consider other penalty operators and more data-driven ways of choosing the smoothing parameter, including links with adaptive or hierarchical shrinkage priors in larger mixed-frequency models.

References

Altissimo, Filippo, Riccardo Cristadoro, Mario Forni, Marco Lippi, and Giovanni Veronese (2010) "New Eurocoin: Tracking Economic Growth in Real Time," *The Review of Economics and Statistics*, 92 (4), 1024–1034, 10.1162/REST_a_00045.

Andreou, Elena, Eric Ghysels, and Andros Kourtellos (2010) "Regression models with mixed sampling frequencies," *Journal of Econometrics*, 158, 246–261, 10.1016/j.jeconom.2010.01.004.

Babii, Andrii, Eric Ghysels, and Jonas Striaukas (2022) "Machine Learning Time Series Regressions with an Application to Nowcasting," *Journal of Business & Economic Statistics*, 40 (3), 1094–1106, 10.1080/07350015.2021.1899933.

Bickel, Peter J and Bas JK Kleijn (2012) "The semiparametric Bernstein–von Mises theorem."

Castillo, I. and R. Nickl (2014) *Bayesian Nonparametric Inference*, 52 of Cambridge Series in Statistical and Probabilistic Mathematics: Cambridge University Press, 10.1017/CBO9781107337862.

Chan, Joshua CC, Aubrey Poon, and Dan Zhu (2025) "Time-Varying Parameter MIDAS Models: Application to Nowcasting US Real GDP," *Journal of Econometrics*, forthcoming.

Chen, Xiaohong and Timothy M. Christensen (2015) "Optimal Uniform Convergence Rates and Asymptotic Normality for Series Estimators Under Weak Dependence and Weak Conditions," *Journal of Econometrics*, 188 (2), 447–465.

Claeskens, Gerda, Tatyana Krivobokova, and Jean D. Opsomer (2009) "Asymptotic Properties of Penalized Spline Estimators," *Biometrika*, 96 (3), 529–544.

Doukhan, Paul (1994) *Mixing: Properties and Examples*, 85 of Lecture Notes in Statistics, New York: Springer.

Doz, Catherine, Domenico Giannone, and Lucrezia Reichlin (2012) "A Quasi–Maximum Likelihood Approach for Large, Approximate Dynamic Factor Models," *The Review of Economics and Statistics*, 94 (4), 1014–1024, 10.1162/REST_a_00225.

Eilers, Paul H. C. and Brian D. Marx (1996) "Flexible smoothing with B-splines and penalties," *Statistical Science*, 11 (2), 89–121.

Foroni, Claudia, Massimiliano Marcellino, and Christian Schumacher (2015) "Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials," *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 178 (1), 57–82, 10.1111/rssa.12043.

Ghosh, Anisha and Oliver Linton (2023) "Estimation with mixed data frequencies: A bias-correction approach," *Journal of Empirical Finance*, 74, 10141, 10.1016/j.jempfin.2023.02.004.

Ghysels, Eric, Pedro Santa-Clara, and Rossen Valkanov (2004) "The MIDAS Touch: Mixed Data Sampling Regression Models," Working paper.

Ghysels, Eric, Arthur Sinko, and Rossen Valkanov (2007) "MIDAS Regressions: Further Results and New Directions," *Econometric Reviews*, 26 (1), 53–90, 10.1080/07474930600972467.

Giannone, Domenico, Lucrezia Reichlin, and David Small (2008) "Nowcasting: The Real-Time Informational Content of Macroeconomic Data," *Journal of Monetary Economics*, 55 (4), 665–676, 10.1016/j.jmoneco.2008.05.010.

Hastie, Trevor, Andrea Montanari, Saharon Rosset, and Ryan J. Tibshirani (2022) "Surprises in High-Dimensional Ridgeless Least Squares Interpolation," *The Annals of Statistics*, 50 (2), 949–986.

Jurado, Kyle, Sydney C. Ludvigson, and Serena Ng (2015) "Measuring Uncertainty," *American Economic Review*, 105 (3), 1177–1216, 10.1257/aer.20131193.

Koop, Gary and Dale J. Poirier (2004) "Bayesian Variants of Some Classical Semiparametric Regression Techniques," *Journal of Econometrics*, 123 (2), 259–282, 10.1016/j.jeconom.2003.12.008.

Koop, Gary, Dale J. Poirier, and Justin L. Tobias (2005) "Semiparametric Bayesian Inference in Multiple Equation Models," *Journal of Applied Econometrics*, 20 (6), 723–747, 10.1002/jae.810.

MacCarthy, Martin and Stephen Snudden (2024) *Predictable by Construction: Assessing Forecast Directional Accuracy of Temporal Aggregates*: LCERPA, Laurier Centre for Economic Research & Policy Analysis.

Mogliani, Matteo and Anna Simoni (2021) "Bayesian MIDAS penalized regressions: Estimation, selection, and prediction," *Journal of Econometrics*, 222 (1), 833–860.

Newey, Whitney K. (1997) "Convergence Rates and Asymptotic Normality for Series Estimators," *Journal of Econometrics*, 79 (1), 147–168.

Newey, Whitney K and Daniel McFadden (1994) "Large sample estimation and hypothesis testing," *Handbook of econometrics*, 4, 2111–2245.

Pettenuzzo, Davide, Riccardo Sabbatucci, and Allan Timmermann (2024) "Financial Statements and Macroeconomic Dynamics," *Working Paper*.

Rossi, Barbara and Tatevik Sekhposyan (2015) "Macroeconomic Uncertainty Indices Based on Nowcast and Forecast Error Distributions," *American Economic Review: Papers & Proceedings*, 105 (5), 650–655, 10.1257/aer.p20151124.

Rudin, Walter (1991) *Functional Analysis*, New York: McGraw–Hill, 2nd edition.

van der Vaart, A. W. (1998) *Asymptotic Statistics*: Cambridge University Press.

Vershynin, Roman (2018) *High-Dimensional Probability: An Introduction with Applications in Data Science*, Cambridge Series in Statistical and Probabilistic Mathematics: Cambridge University Press, 10.1017/9781108231596.

Wang, Hansheng and Chenlei Leng (2008) "A note on adaptive group lasso," *Computational Statistics and Data Analysis*, 52 (12), 5277–5286.

White, Halbert (2001) *Asymptotic Theory for Econometricians*, San Diego: Academic Press, revised edition.

Online Appendix

A Proof of Theorem 1

This appendix provides a proof to Theorem 1 in the main text.

We shall first present some useful lemmas.

Lemma 1. *By Assumption 1 and the assumptions in Theorem 1, we have*

1. $\{(Z_t(\cdot), u_t) : t \in \mathbb{Z}\}$ *is strictly stationary.*
2. *Let* $\alpha_{(Z,u)}(k)$ *denote the strong mixing coefficients of* $\{(Z_t(\cdot), u_t)\}$*. Then*

$$\alpha_{(Z,u)}(k) \;\le\; \alpha_Z(k) + \alpha_u(k), \qquad \alpha_Z(k) = 0 \;\; \textit{for all } k > \tau,$$

   so in particular $\sum_{k\ge 1} \alpha_{(Z,u)}(k)^{\delta/(2+\delta)} < \infty$*.*

3. *For every* $f \in H$*,*

$$E\big[\,|Z_t(f)|^{2+\delta}\,\big] \;\le\; C_{2+\delta}\,\|f\|_H^{2+\delta} \;<\; \infty,$$

   with a constant $C_{2+\delta}$ *independent of* t *and* f*.*

Proof. Write

$$Z_t(f) = D(f) + M_t(f) \quad \text{with} \quad D(f) := \int_0^\tau f(s)^\top \boldsymbol{\mu}\, ds, \qquad M_t(f) := \int_0^\tau f(s)^\top \boldsymbol{\Sigma}\, d\mathbf{W}(t-s).$$

Since $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ are constant and Brownian increments are stationary and the deterministic integrands are time shifts of the same functions in H. Thus $\{Z_t(\cdot)\}$ is strictly stationary. Because (u_t) is strictly stationary and independent of $\mathbf{W}$, the pair $\{(Z_t(\cdot), u_t)\}$ is strictly stationary. By a change of variables,

$$M_t(f) = \int_{t-\tau}^{t} \psi_f(r)\, d\mathbf{W}(r), \qquad \psi_f(r) := \Sigma^\top f(t-r),$$

so for each fixed t the random variables $\{M_t(f) : f \in H\}$ are measurable with respect to

$$\mathcal{G}_t := \sigma\big(\mathbf{W}(r) - \mathbf{W}(t-\tau) : r \in [t-\tau, t]\big).$$

If $k > \tau$, then the intervals $[t-\tau, t]$ and $[t+k-\tau, t+k]$ are disjoint. By the independent increments property of Brownian motion, $\mathcal{G}_t$ and $\mathcal{G}_{t+k}$ are independent, hence

$$\alpha_Z(k) = \alpha_M(k) = 0 \qquad \text{for all } k > \tau.$$

For the joint process $\{(Z_t(\cdot), u_t)\}$, since (u_t) is independent of $\mathbf{W}$ and has strong mixing coefficients $\alpha_u(k)$, the closure property of strong mixing (Ibragimov-Linnik) yields

$$\alpha_{(Z,u)}(k) \le \alpha_Z(k) + \alpha_u(k).$$

Because $\alpha_Z(k) = 0$ for $k > \tau$ and $\sum_{k\ge 1} \alpha_u(k)^{\delta/(2+\delta)} < \infty$ by assumption, we conclude that

$$\sum_{k\ge 1} \alpha_{(Z,u)}(k)^{\delta/(2+\delta)} < \infty.$$

By Cauchy-Schwarz,

$$|D(f)| \;=\; \Big| \int_0^\tau f(s)^\top \boldsymbol{\mu}\, ds \Big| \;\le\; \|f\|_H \Big(\int_0^\tau \|\boldsymbol{\mu}\|^2\, ds \Big)^{1/2} \;=\; \sqrt{\tau}\, \|\boldsymbol{\mu}\|\, \|f\|_H.$$

For the martingale term,

$$V(f) \;=\; \int_0^\tau f(s)^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top f(s)\, ds \;=\; \int_0^\tau \langle f(s), (\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top) f(s) \rangle\, ds.$$

Since $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top$ is a fixed symmetric positive definite matrix,

$$V(f) \;\le\; \lambda_{\max}(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top)\, \|f\|_H^2,$$

where $\lambda_{\max}$ is the largest eigenvalue. Hence, for any $\delta > 0$, the Gaussian moment formula gives

$$E|M_t(f)|^{2+\delta} = c_{2+\delta}\, V(f)^{1+\delta/2} \;\le\; C_1\, \|f\|_H^{2+\delta},$$

with C_1 depending only on δ and $\boldsymbol{\Sigma}$. Combining this with the bound for $D(f)$ and using $(a+b)^{2+\delta} \le 2^{1+\delta}(a^{2+\delta} + b^{2+\delta})$,

$$E|Z_t(f)|^{2+\delta} \;\le\; C_{2+\delta}\, \|f\|_H^{2+\delta},$$

for a constant $C_{2+\delta}$ depending only on δ, τ, $\boldsymbol{\mu}$, and $\boldsymbol{\Sigma}$, but not on t or f. □

Lemma 2. *Fix $p = 4 + \delta > 4$. For $h = \tau/m$ define the bin increments*

$$\Delta \mathbf{X}_{t,\ell} := \mathbf{X}\big(t - (\ell-1)h\big) - \mathbf{X}\big(t - \ell h\big) = \boldsymbol{\mu}\, h + \boldsymbol{\Sigma}\, \Delta \mathbf{W}_{t,\ell}, \qquad \ell = 1, \dots, m,$$

where $\Delta \mathbf{W}_{t,\ell} \sim \mathcal{N}(0, h I_J)$ are independent across ℓ and independent of the past. Let $\mathbf{X}_t^{(m)} := (\Delta \mathbf{X}_{t,1}^\top, \dots, \Delta \mathbf{X}_{t,m}^\top)^\top \in \mathbf{R}^{mJ}$. Then there exists a constant $C_p < \infty$ such that, uniformly in t and m,

$$\mathbb{E}\big[\|\mathbf{X}_t^{(m)}\|^p\big] \;\le\; C_p\Big(\tau^{p/2} + \tau^p\Big) \;<\; \infty.$$

In particular, $\mathbb{E}\big[\|\mathbf{X}_t^{(m)}\|^{4+\delta}\big] < \infty$ uniformly in m.

Proof. Each increment is Gaussian: $\Delta \mathbf{X}_{t,\ell} \sim \mathcal{N}(\boldsymbol{\mu} h,\, A h)$, $A = \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top$. By $(a+b)^p \le 2^{p-1}(a^p + b^p)$,

$$\mathbb{E}\|\Delta \mathbf{X}_{t,\ell}\|^p \;\le\; 2^{p-1}\Big(\|\boldsymbol{\mu}\|^p h^p \;+\; \mathbb{E}\|\boldsymbol{\Sigma}\, \Delta \mathbf{W}_{t,\ell}\|^p\Big).$$

Since $\Delta \mathbf{W}_{t,\ell} \overset{d}{=} \sqrt{h}\, \mathbf{Z}$ with $\mathbf{Z} \sim \mathcal{N}(0, I_J)$,

$$\mathbb{E}\|\boldsymbol{\Sigma}\, \Delta \mathbf{W}_{t,\ell}\|^p = h^{p/2}\, \mathbb{E}\|\boldsymbol{\Sigma}\, \mathbf{Z}\|^p \;\le\; C_{p,J}\, (\operatorname{tr} A)^{p/2}\, h^{p/2},$$

where $C_{p,J} < \infty$ depends only on p and J.[3] Hence

$$\mathbb{E}\|\Delta \mathbf{X}_{t,\ell}\|^p \;\le\; C_p\big(h^p + h^{p/2}\big),$$

for a constant C_p depending on $p, J, \|\boldsymbol{\mu}\|$, and $\operatorname{tr} A$ (but not on m).

For $p \ge 2$ and nonnegative a_ℓ, $\big(\sum_{\ell=1}^m a_\ell^2\big)^{p/2} \le m^{\frac{p}{2}-1} \sum_{\ell=1}^m a_\ell^p$, so

$$\mathbb{E}\|\mathbf{X}_t^{(m)}\|^p = \mathbb{E}\Big(\sum_{\ell=1}^m \|\Delta \mathbf{X}_{t,\ell}\|^2\Big)^{p/2} \le m^{\frac{p}{2}-1} \sum_{\ell=1}^m \mathbb{E}\|\Delta \mathbf{X}_{t,\ell}\|^p \le m^{\frac{p}{2}}\, C_p\big(h^p + h^{p/2}\big).$$

With $h = \tau/m$, $m^{\frac{p}{2}} h^p = \tau^p m^{-p/2} \le \tau^p$ and$m^{\frac{p}{2}} h^{p/2} = \tau^{p/2}$. Therefore

$$\mathbb{E}\|\mathbf{X}_t^{(m)}\|^p \;\le\; C_p\big(\tau^{p/2} + \tau^p\big),$$

uniformly in m. □

Lemma 3. *Let $A := \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top$. Then the bilinear form*

$$q(f, g) \;:=\; E[\, Z_t(f)\, Z_t(g)\,]$$

[3]For Gaussian vectors, $\mathbb{E}\|A^{1/2}\mathbf{Z}\|^p \le C_{p,J}(\operatorname{tr} A)^{p/2}$ follows from rotational invariance and polynomial equivalence of moments in finite dimension.

admits the representation

$$q(f,g) = \left(\int_0^\tau f(s)\,ds\right)^\top \boldsymbol{\mu}\boldsymbol{\mu}^\top \left(\int_0^\tau g(s)\,ds\right) \;+\; \int_0^\tau f(s)^\top A\, g(s)\,ds.$$

Consequently, there exists a bounded, self-adjoint, positive operator $Q : H \to H$ such that $q(f,g) = \langle f, Qg\rangle_H$ for all $f, g \in H$, and Q is given explicitly by

$$(Qg)(s) \;=\; A\,g(s) \;+\; \boldsymbol{\mu}\boldsymbol{\mu}^\top \int_0^\tau g(u)\,du, \qquad s \in [0,\tau] \;\; a.e.$$

Moreover,

$$0 < c \;\le\; \frac{\langle f, Qf\rangle_H}{\|f\|_H^2} \;\le\; C < \infty \qquad \textit{for all } f \in H,$$

with $c = \lambda_{\min}(A)$ (in particular, $c > 0$ if A is positive definite) and $C = \lambda_{\max}(A) + \tau\|\boldsymbol{\mu}\|^2$.

Proof. Since $D(f)$ is deterministic and $E[M_t(f)] = 0$,

$$q(f,g) = D(f)D(g) + E\big[M_t(f)M_t(g)\big].$$

The drift part equals

$$D(f)D(g) = \Big(\int_0^\tau f\Big)^\top \boldsymbol{\mu}\boldsymbol{\mu}^\top \Big(\int_0^\tau g\Big).$$

For the martingale part, by Itô isometry with constant integrand matrix,

$$E\big[M_t(f)M_t(g)\big] = \int_0^\tau f(s)^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top g(s)\,ds = \int_0^\tau f(s)^\top A\, g(s)\,ds.$$

Adding the two terms gives the displayed formula for $q(f,g)$.

Then

$$\langle f, Qg\rangle_H = \int_0^\tau f(s)^\top A g(s)\,ds + \left(\int_0^\tau f(s)\,ds\right)^\top \boldsymbol{\mu}\boldsymbol{\mu}^\top \left(\int_0^\tau g(u)\,du\right) = q(f,g),$$

so $q(\cdot,\cdot)$ is represented by Q. Self-adjointness is immediate since A is symmetric and the rank-one term is symmetric. Positivity follows from

$$\langle f, Qf\rangle_H = \int_0^\tau f(s)^\top A f(s)\,ds + \left\|\boldsymbol{\mu}^\top \int_0^\tau f(s)\,ds\right\|^2 \;\ge\; 0.$$

For the bounds, first the upper bound: by Cauchy-Schwarz,

$$\Big|\boldsymbol{\mu}^\top \int_0^\tau f(s)\,ds\Big|^2 \le \|\boldsymbol{\mu}\|^2 \Big\| \int_0^\tau f(s)\,ds\Big\|^2 \le \|\boldsymbol{\mu}\|^2\,\tau\,\|f\|_H^2,$$

and $\int f^\top A f ds \le \lambda_{\max}(A)\|f\|_H^2$, hence

$$\langle f, Qf\rangle_H \le \big(\lambda_{\max}(A) + \tau\|\boldsymbol{\mu}\|^2\big)\,\|f\|_H^2.$$

For the lower bound, if A is positive definite then

$$\langle f, Qf\rangle_H \;\ge\; \int_0^\tau f(s)^\top A f(s)\,ds \;\ge\; \lambda_{\min}(A)\,\|f\|_H^2,$$

so $c = \lambda_{\min}(A) > 0$. □

Before stating the next lemma, note that under the stated α-mixing condition and a finite $(2+\delta)$ moment for u_t, the autocovariances are absolutely summable, i.e. $\sum_{k\in\mathbb{Z}} |\gamma_u(k)| < \infty$.

Lemma 4. *Consider the bilinear form*

$$\omega(f,g) \;:=\; \sum_{k\in\mathbb{Z}} E\big[\, Z_t(f)u_t\; Z_{t-k}(g)u_{t-k}\,\big].$$

Then the series converges absolutely and defines a bounded, self-adjoint, positive operator $\Omega : H \to H$ via $\omega(f,g) = \langle f, \Omega g\rangle_H$, such that

$$\langle f, \Omega g\rangle_H = \Big(\sum_{k\in\mathbb{Z}} \gamma_u(k)\Big)\Big(\int_0^\tau f(s)^\top \mu\,ds\Big)\Big(\int_0^\tau g(s)^\top \mu\,ds\Big) + \sum_{|k|\le\tau} \gamma_u(k)\, I_k(f,g),$$

where the overlap integrals $I_k(f,g)$ are given by

$$I_k(f,g) = \begin{cases} \displaystyle\int_k^\tau f(s)^\top(\Sigma\Sigma^\top)\, g(s-k)\,ds, & k \ge 0, \\[2ex] \displaystyle\int_0^{\tau+k} f(s)^\top(\Sigma\Sigma^\top)\, g(s-k)\,ds, & k < 0. \end{cases}$$

Proof. Because u_t is independent of $\mathbf{W}$, for every $k \in \mathbb{Z}$,

$$E\big[Z_t(f)u_t\, Z_{t-k}(g)u_{t-k}\big] = E\big[Z_t(f)Z_{t-k}(g)\big]\; E[u_t u_{t-k}] = \gamma_u(k)\, E\big[Z_t(f)Z_{t-k}(g)\big].$$

Write $Z_t(f) = D(f) + M_t(f)$ with $E[M_t(f)] = 0$. Then

$$E\big[Z_t(f)Z_{t-k}(g)\big] = D(f)D(g) \;+\; E\big[M_t(f)M_{t-k}(g)\big].$$

By Itô isometry for vector integrals,

$$\begin{aligned} E[M_t(f)M_{t-k}(g)] &= \int_{\mathbb{R}} f(t-r)^\top A\, g(t-k-r)\, \mathbf{1}_{[t-\tau,\,t]}(r)\, \mathbf{1}_{[t-k-\tau,\,t-k]}(r)\, dr \\ &= \int_k^\tau f(s)^\top A\, g(s-k)\, ds. \end{aligned}$$

and the expression is 0 when $k > \tau$ (no overlap).Similarly, for $k < 0$, $E[M_t(f)M_{t-k}(g)] = \int_0^{\tau+k} f(s)^\top A\, g(s-k)\, ds$. Therefore, form

$$\omega(f,g) = \sum_{k\in\mathbb{Z}} \gamma_u(k)\, D(f)D(g) \;+\; \sum_{|k|\le\tau} \gamma_u(k) \begin{cases} \int_k^\tau f(s)^\top A\, g(s-k)\, ds, & k \ge 0, \\ \int_0^{\tau+k} f(s)^\top A\, g(s-k)\, ds, & k < 0, \end{cases}$$

with $\gamma_u(k) = E[u_t u_{t-k}]$.

By Cauchy-Schwarz and $\|h_\mu\|_H^2 = \int_0^\tau \|\mu\|^2\, ds = \tau\|\mu\|^2$,

$$|D(f)D(g)| = \big|\langle f, h_\mu\rangle_H \langle h_\mu, g\rangle_H\big| \le \|f\|_H \|h_\mu\|_H \cdot \|g\|_H \|h_\mu\|_H = \tau\|\mu\|^2 \|f\|_H \|g\|_H.$$

For the overlap terms, note that for each k with $|k| \le \tau$,

$$\Big|\int f(s)^\top A\, g(s-k)\, ds\Big| \le \lambda_{\max}(A)\, \|f\|_H \|g\|_H,$$

since A is constant symmetric positive semi definite. Therefore

$$\sum_{k\in\mathbb{Z}} |\gamma_u(k)|\, |D(f)D(g)| \;\le\; \Big(\sum_{k\in\mathbb{Z}} |\gamma_u(k)|\Big)\, \tau\|\mu\|^2\, \|f\|_H \|g\|_H,$$

and

$$\sum_{|k|\le\tau} |\gamma_u(k)| \Big|\int f(s)^\top A\, g(s-k)\, ds\Big| \;\le\; \Big(\sum_{|k|\le\tau} |\gamma_u(k)|\Big)\, \lambda_{\max}(A)\, \|f\|_H \|g\|_H.$$

Since $\sum_k |\gamma_u(k)| < \infty$, the bilinear form $\omega(f,g)$ converges absolutely and

$$|\omega(f,g)| \;\le\; C\, \|f\|_H \|g\|_H$$

for some finite C depending only on τ, μ, Σ and the $\gamma_u(k)$. Hence the induced operator Ω is bounded.

Both the drift term and the overlap integrals are symmetric in (f, g) (since $A = \Sigma\Sigma^\top$ is symmetric and $\gamma_u(k) = \gamma_u(-k)$). Hence $\omega(f, g) = \omega(g, f)$, so Ω is self-adjoint.

It is symmetric by construction and nonnegative on the diagonal, so the induced operator Ω is bounded, self-adjoint, and positive. □

We now prove Theorem 1.

Proof. By the data-generating relation $y_t = Z_t(\beta) + u_t$ with $\mathbb{E}[u_t \mid \mathcal{F}_{t^-}] = 0$,

$$\begin{aligned} L(f) - L(\beta) &= \mathbb{E}\big[(y_t - Z_t(f))^2\big] - \mathbb{E}\big[(y_t - Z_t(\beta))^2\big] \\ &= \mathbb{E}\big[(Z_t(\beta) + u_t - Z_t(f))^2\big] - \mathbb{E}\big[u_t^2\big] \\ &= \mathbb{E}\big[(Z_t(\beta - f))^2\big] \\ &=: q(\beta - f, \beta - f) = \langle \beta - f,\ Q(\beta - f)\rangle_H \end{aligned}$$

Hence $L(\cdot)$ is strictly convex and its unique minimizer over H_m is the Q-orthogonal projection

$$\beta_m \ := \ \arg\min_{f \in H_m} L(f).$$

By construction, $\beta_m \in H_m$ solves

$$\beta_m \ = \ \arg\min_{g \in H_m} L(g) \quad \Longleftrightarrow \quad \exists\, \theta_m^0 \in \mathbb{R}^{mJ} \text{ such that } \beta_m = \Pi_m \theta_m^0 = \arg\min_{\theta \in \mathbb{R}^{mJ}} L(\Pi_m \theta).$$

Since the union of step-function spaces $\bigcup_m H_m$ is dense in $H = L^2([0,1]; \mathbb{R}^J)$, it follows that for any $\beta \in H$ and any $\varepsilon > 0$, there exist m and $\tilde{\theta}_m \in \mathbb{R}^{mJ}$ such that

$$\|\Pi_m \tilde{\theta}_m - \beta\|_H < \varepsilon.$$

By minimality of β_m over H_m,

$$0 \ \le \ L(\beta_m) - L(\beta) \ \le \ L(\Pi_m \tilde{\theta}_m) - L(\beta) = \langle \beta - \Pi_m \tilde{\theta}_m,\ Q(\beta - \Pi_m \tilde{\theta}_m)\rangle_H \ \le \ C\, \varepsilon^2,$$

using $L(g) - L(\beta) = \langle \beta - g, Q(\beta - g)\rangle_H$ and the upper bound $\langle h, Qh\rangle_H \le C\|h\|_H^2$. The lower bound $\langle h, Qh\rangle_H \ge c\|h\|_H^2$ then gives

$$c\, \|\beta_m - \beta\|_H^2 \ \le \ L(\beta_m) - L(\beta) \ \le \ C\, \varepsilon^2.$$

As $m \to \infty$ we can take $\varepsilon \to 0$, hence $\|\beta_m - \beta\|_H \to 0$.

For $f \in H_m$ we can write $f = \Pi_m \theta$ with $\theta = (\theta_1, \dots, \theta_m) \in \mathbb{R}^{mJ}$, where $f(s) = \theta_\ell$ on bin $I_\ell = \big((\ell-1)\frac{\tau}{m},\, \ell\frac{\tau}{m}\big]$. Then

$$\|f\|_H^2 = \int_0^\tau \|f(s)\|^2\, ds = \sum_{\ell=1}^m \int_{I_\ell} \|\theta_\ell\|^2\, ds = \frac{\tau}{m} \sum_{\ell=1}^m \|\theta_\ell\|^2 = \frac{\tau}{m}\, \|\theta\|^2,$$

where $\|\theta\|^2 := \sum_{\ell=1}^m \|\theta_\ell\|^2$. Hence the ball of radius B in H restricted to H_m is

$$H_m(B) := \{f \in H_m : \|f\|_H \le B\} = \big\{\Pi_m \theta : \|\theta\| \le B\sqrt{m/\tau}\big\}.$$

Now

$$Z_t(f) := \int_0^\tau f(s)\, dX(t-s) = \sum_{\ell=1}^m \theta_\ell^\top \Big(X\big(t - (\ell-1)\tfrac{\tau}{m}\big) - X\big(t - \ell\tfrac{\tau}{m}\big)\Big).$$

Let $\Delta X_{t,\ell} := X(t - (\ell-1)\frac{\tau}{m}) - X(t - \ell\frac{\tau}{m})$ and $X_t^{(m)} := (\Delta X_{t,1}^\top, \dots, \Delta X_{t,m}^\top)^\top$. Then

$$Z_t(f) = \theta^\top X_t^{(m)}.$$

Define the sample and population Gram operators $\widehat{Q}_{T,m}, Q_m : H_m \to H_m$ and the score elements $\widehat{\gamma}_{T,m}, \gamma_m \in H_m$ by

$$\langle f, \widehat{Q}_{T,m} g\rangle_H := \frac{1}{T}\sum_{t=1}^T Z_t(f)\, Z_t(g), \qquad \langle f, Q_m g\rangle_H := \mathbb{E}[Z_t(f)\, Z_t(g)], \quad f, g \in H_m,$$

and

$$\langle f, \widehat{\gamma}_{T,m}\rangle_H := \frac{1}{T}\sum_{t=1}^T y_t\, Z_t(f), \qquad \langle f, \gamma_m\rangle_H := \mathbb{E}[y_t\, Z_t(f)], \quad f \in H_m.$$

The OLS estimator $\widehat{\beta}_{T,m} \in H_m$ then satisfies the normal equations

$$\widehat{Q}_{T,m}\, \widehat{\beta}_{T,m} \;=\; \widehat{\gamma}_{T,m}.$$

By the law of large numbers for α-mixing sequences (see e.g. White (2001), Thm. 3.49) each entry of $\widehat{Q}_{T,m}$ and $\widehat{\gamma}_{T,m}$ converges uniformly to its population counterpart. Moreover, for any symmetric $m \times m$ matrix $K = (K_{jk})$,

$$\|K\|_{\mathrm{op}} \;\le\; \|K\|_F \;\le\; m \max_{j,k} |K_{jk}|,$$

so that uniform entrywise convergence implies convergence in operator norm, provided $m =$

$o(T)$. In particular,

$$\left\|\widehat{Q}_{T,m} - Q_m\right\|_{\text{op}} \xrightarrow{p} 0, \qquad \left\|\widehat{\gamma}_{T,m} - \gamma_m\right\| \xrightarrow{p} 0,$$

as $T \to \infty$, uniformly over $m = m_T = o(T)$.

From Lemma 2, $\langle h, Q_m h\rangle_H \geq c\|h\|_H^2$ for all $h \in H_m$. By the uniform LLN, $\|\widehat{Q}_{T,m} - Q_m\|_{\text{op}} \to_p 0$, so for any $\varepsilon \in (0, c)$, with probability $\to 1$,

$$\langle h, \widehat{Q}_{T,m} h\rangle_H \;\geq\; (c - \varepsilon)\|h\|_H^2 \qquad \forall\, h \in H_m.$$

Expanding $S_{T,m}(f)$,

$$S_{T,m}(f) = \langle f, \widehat{Q}_{T,m} f\rangle_H - 2\langle f, \widehat{\gamma}_{T,m}\rangle_H + \tfrac{1}{T}\sum_{t=1}^{T} u_t^2 + \text{const.}$$

By Cauchy-Schwarz and the LLN for $\widehat{\gamma}_{T,m}$ and $T^{-1}\sum u_t^2$, there exists $B < \infty$ (nonrandom) such that, w.p.a.1,

$$S_{T,m}(f) \;\geq\; (c - \varepsilon)\|f\|_H^2 - o_p(1)\,\|f\|_H - O_p(1).$$

Hence $S_{T,m}(f) \to \infty$ as $\|f\|_H \to \infty$ uniformly on H_m w.p.a.1. Therefore the empirical minimizer $\widehat{f}_{T,m}$ lies in some ball $H_m(B)$ w.p.a.1.

Now

$$S_{T,m}(f) - L(f) = (I) - 2(II) + (III),$$

with

$$(I) = \frac{1}{T}\sum_{t=1}^{T} Z_t(f)^2 - \mathbb{E}[Z_t(f)^2], \quad (II) = \frac{1}{T}\sum_{t=1}^{T} u_t Z_t(f) - \mathbb{E}[u_t Z_t(f)], \quad (III) = \frac{1}{T}\sum_{t=1}^{T} u_t^2 - \mathbb{E}[u_t^2].$$

Term (III) is scalar and converges to zero by the LLN for α-mixing sequences ((White, 2001, Thm. 3.49)).

For the first two terms, note that for $f = \Pi_m \theta \in H_m$, we have

$$(I) = \langle f, (\widehat{Q}_{T,m} - Q_m) f\rangle_H, \qquad (II) = \langle f, \widehat{\gamma}_{T,m} - \gamma_m\rangle_H.$$

Hence, uniformly over $\|\theta\| \le R_T$ (equivalently, $\|f\|_H \le B$),

$$\sup_{\|\theta\|\le R_T} |(I)| \le R_T^2 \|\widehat{Q}_{T,m} - Q_m\|_{\mathrm{op}}, \qquad \sup_{\|\theta\|\le R_T} |(II)| \le R_T \|\widehat{\gamma}_{T,m} - \gamma_m\|_H,$$

where $R_T = B\sqrt{m/\tau}$. Each entry of the matrix representation of $\widehat{Q}_{T,m}$ and the vector representation of $\widehat{\gamma}_{T,m}$ converges by the LLN for α-mixing arrays. To make this uniform over θ, we proceed via an ε-net. For any $f = \Pi_m\theta$ and $f' = \Pi_m\theta'$,

$$\begin{aligned}
\|Z_\theta - Z_{\theta'}\|_2 &\le \|\theta - \theta'\| \, (\mathbb{E}\|X_t^{(m)}\|^2)^{1/2}, \\
\|Z_\theta^2 - Z_{\theta'}^2\|_2 &\le 2R_T (\mathbb{E}\|X_t^{(m)}\|^4)^{1/2} \|\theta - \theta'\|, \\
\|u_t Z_\theta - u_t Z_{\theta'}\|_2 &\le \|\theta - \theta'\| \, (\mathbb{E}[u_t^2 \|X_t^{(m)}\|^2])^{1/2}.
\end{aligned}$$

Thus the maps $\theta \mapsto Z_\theta^2$ and $\theta \mapsto u_t Z_\theta$ are Lipschitz in $L^2(P)$. Cover the Euclidean ball $B_2^d(R_T)$ by an ε-net of size $N(\varepsilon, B_2^d(R_T)) \le \left(\frac{3R_T}{\varepsilon}\right)^d$. That is: for every $\theta \in B_2^d(R_T)$ there exists $\tilde{\theta} \in \mathcal{N}_\varepsilon$ with $\|\theta - \tilde{\theta}\| \le \varepsilon$, see Vershynin (2018) . At each fixed $\tilde{\theta} \in \mathcal{N}_\varepsilon$, the quantities $(I)(\tilde{\theta}), (II)(\tilde{\theta})$ are averages of stationary α-mixing sequences with finite variance from the previous lemma. Therefore, by Chebyshev's inequality,

$$\Pr\big(|(I)(\tilde{\theta})| > \delta\big) \;\le\; \frac{C}{T\delta^2}, \qquad \Pr\big(|(II)(\tilde{\theta})| > \delta\big) \;\le\; \frac{C'}{T\delta^2},$$

for constants C, C' independent of T. Now take a union bound over all net points:

$$\Pr\Big(\max_{\tilde{\theta}\in\mathcal{N}_\varepsilon} \big(|(I)(\tilde{\theta})| \vee |(II)(\tilde{\theta})| > \delta\big)\Big) \;\le\; |\mathcal{N}_\varepsilon| \cdot \frac{C''}{T\delta^2}.$$

Since $|\mathcal{N}_\varepsilon| \le (3R_T/\varepsilon)^d$, the right-hand side goes to zero whenever

$$\log|\mathcal{N}_\varepsilon| = d \log\Big(\tfrac{3R_T}{\varepsilon}\Big) = o(T).$$

Finally, extend from the net to the full ball. By the Lipschitz bounds established earlier, if $\|\theta - \tilde{\theta}\| \le \varepsilon$ then

$$|(I)(\theta) - (I)(\tilde{\theta})| \le C_1 R_T^2 \varepsilon, \qquad |(II)(\theta) - (II)(\tilde{\theta})| \le C_2 R_T \varepsilon,$$

for constants C_1, C_2 depending only on the moments of $X_t^{(m)}$ and u_t. Thus

$$\sup_{\|\theta\|\le R_T} |(I)(\theta)| \le \max_{\tilde{\theta}\in\mathcal{N}_\varepsilon} |(I)(\tilde{\theta})| + C_1 R_T^2 \varepsilon,$$

and similarly for (II). Letting $T \to \infty$ and then $\varepsilon \to 0$ shows that

$$\sup_{\|\theta\| \le R_T} |(I)| = o_p(1), \qquad \sup_{\|\theta\| \le R_T} |(II)| = o_p(1).$$

Now, we have

$$d \log\left(\tfrac{R_T}{\varepsilon}\right) = o(T).$$

Since $d = mJ$ and $R_T^2 \asymp m$, the condition is

$$m \log T = o(T),$$

equivalently $m = o(T/\log T)$. Together with $(III) \to_p 0$, we obtain

$$\sup_{f \in H_m(B)} |S_{T,m}(f) - L(f)| \xrightarrow{p} 0,$$

establishing uniform convergence of the sample criterion.

By uniform convergence of $S_{T,m}(\cdot)$ to $L(\cdot)$ on $H_m(B)$ and uniqueness of the minimizer β_m, the argmin continuity theorem (Theorem 2.7 in Newey and McFadden (1994)) gives

$$\|\widehat{f}_{T,m} - \beta_m\|_H \xrightarrow{p} 0, \qquad T \to \infty,\ m = o(T).$$

Next, decompose

$$\|\Pi_m \widehat{\beta}_{T,m} - \beta\|_H \ \le\ \|\widehat{f}_{T,m} - \beta_m\|_H \ +\ \|\beta_m - \beta\|_H.$$

The first term converges to zero in probability by the argmin result. The second term converges to zero deterministically as $m \to \infty$ (since β_m is the projection of β onto H_m and $\bigcup_m H_m$ is dense in H). Therefore,

$$\|\Pi_m \widehat{\beta}_{T,m} - \beta\|_H \xrightarrow{p} 0, \qquad T \to \infty,\ m = m_T \to \infty,\ m_T = o(T).$$

This proves consistency in H.

Define the score map $\xi_t : H \to \mathbb{R}$ by

$$\xi_t(f) := u_t Z_t(f), \qquad f \in H, \quad t \in \mathbb{Z},$$

that is zero mean, square integrable. Finally, by assumption the joint process $\{(Z_t(\cdot), u_t)\}$ is strictly stationary and α-mixing with $\sum_{k \ge 1} \alpha(k)^{\delta/(2+\delta)} < \infty$ for some $\delta > 0$, $\xi_t(f) = u_t Z_t(f)$

is just the product of the coordinates of this process, $\{\xi_t(f)\}$ inherits strict stationarity and α-mixing with the same coefficients. At the population level, the score element $\gamma \in H$ is defined by

$$\langle f, \gamma \rangle_H = \mathbb{E}[y_t Z_t(f)], \qquad f \in H.$$

Since $y_t = \langle \beta, Z_t(\cdot) \rangle_H + u_t$, we have

$$\langle f, \gamma \rangle_H = \mathbb{E}\big[\big(\langle \beta, Z_t(\cdot) \rangle_H + u_t\big) Z_t(f)\big] = \mathbb{E}[\langle \beta, Z_t(\cdot) \rangle_H \, Z_t(f)] + \mathbb{E}[u_t Z_t(f)].$$

Hence

$$\langle f, \gamma \rangle_H = \mathbb{E}[\langle \beta, Z_t(\cdot) \rangle_H \, Z_t(f)] = \langle f, Q\beta \rangle_H,$$

which shows $\gamma = Q\beta$. From

$$\widehat{Q}_{T,m} \, (\widehat{\beta}_{T,m} - \beta) = -(\widehat{Q}_{T,m} - Q)\beta + (\widehat{\gamma}_{T,m} - \gamma),$$

test any $f \in H_m$

$$\begin{aligned}
\big\langle f, \ -(\widehat{Q}_{T,m} - Q)\beta + (\widehat{\gamma}_{T,m} - \gamma) \big\rangle_H &= -\Big\{ \tfrac{1}{T} \sum_{t=1}^{T} Z_t(f) Z_t(\beta) - \mathbb{E}[Z_t(f) Z_t(\beta)] \Big\} \\
&\quad + \ \Big\{ \tfrac{1}{T} \sum_{t=1}^{T} y_t Z_t(f) - \mathbb{E}[y_t Z_t(f)] \Big\} \\
&= \tfrac{1}{T} \sum_{t=1}^{T} \big(y_t Z_t(f) - Z_t(f) Z_t(\beta) \big) \ - \ \mathbb{E}\big[y_t Z_t(f) - Z_t(f) Z_t(\beta) \big] \\
&= \frac{1}{T} \sum_{t=1}^{T} \xi_t(f).
\end{aligned}$$

By the Riesz-Fréchet representation theorem (Rudin, 1991), there exists a unique $\xi_t \in H_m$ such that $\xi_t(f) = \langle f, \xi_t \rangle_H$ for all $f \in H_m$. Set $a := \widehat{Q}_{T,m}(\widehat{\beta}_{T,m} - \beta)$ and $b := \frac{1}{T} \sum_{t=1}^{T} \xi_t$. From the previous display, $\langle f, a \rangle_H = \langle f, b \rangle_H$ for all $f \in H_m$, hence $\langle f, a - b \rangle_H = 0$ for all $f \in H_m$. Taking $f = a - b \in H_m$ yields $\|a - b\|_H^2 = 0$, so $a = b$. Since each $\xi_t \in H_m$ and $\widehat{Q}_{T,m} : H_m \to H_m$, both sides are in H_m, giving

$$\widehat{Q}_{T,m}(\widehat{\beta}_{T,m} - \beta) = \frac{1}{T} \sum_{t=1}^{T} \xi_t \in H_m.$$

By the curvature of Q and restricting to H_m gives $\langle h, Q_m h \rangle_H \geq c \|h\|_H^2$ for all $h \in H_m$, hence $\lambda_{\min}(Q_m) \geq c$ and Q_m is invertible with $\|Q_m^{-1}\|_{\text{op}} \leq 1/c$. From the previous consistency

proof, fix $\varepsilon \in (0, c)$; with probability tending to 1, $\|\widehat{Q}_{T,m} - Q_m\|_{\text{op}} < \varepsilon$. Then, for any $h \in H_m$,

$$\langle h, \widehat{Q}_{T,m} h\rangle_H = \langle h, Q_m h\rangle_H + \langle h, (\widehat{Q}_{T,m} - Q_m)h\rangle_H \;\geq\; (c - \varepsilon)\,\|h\|_H^2.$$

Taking $\varepsilon = c/2$ yields

$$\langle h, \widehat{Q}_{T,m} h\rangle_H \;\geq\; \tfrac{c}{2}\,\|h\|_H^2 \qquad \forall\, h \in H_m$$

w.p.a.1, so $\widehat{Q}_{T,m}$ is (strictly) positive definite and hence invertible on H_m, with

$$\|(\widehat{Q}_{T,m})^{-1}\|_{\text{op}} \;\leq\; \frac{2}{c} \quad \text{w.p.a.1.}$$

Since $\widehat{\beta}_{T,m} \in H_m$ (so $\Pi_m \widehat{\beta}_{T,m} = \widehat{\beta}_{T,m}$),

$$\sqrt{T}\left(\Pi_m \widehat{\beta}_{T,m} - \beta\right) = (\widehat{Q}_{T,m})^{-1}\,\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \xi_t.$$

Define

$$S_T \;:=\; \frac{1}{\sqrt{T}}\sum_{t=1}^{T} \xi_t \;\in H, \qquad \text{so that} \qquad \langle f, S_T\rangle_H \;=\; \frac{1}{\sqrt{T}}\sum_{t=1}^{T} \xi_t(f) \;\; \forall f \in H.$$

We now derive the asymptotic distribution of fixed linear functionals of $\Pi_m \widehat{\beta}_{T,m} - \beta$. Fix any $\psi \in H$. Taking inner products with ψ yields

$$\sqrt{T}\,\langle \Pi_m \widehat{\beta}_{T,m} - \beta, \psi\rangle_H = \left\langle (\widehat{Q}_{T,m})^{-1} S_T, \psi \right\rangle_H = \left\langle S_T, (\widehat{Q}_{T,m})^{-1} \Pi_m \psi \right\rangle_H.$$

By the operator norm bound established above and the consistency $\|\widehat{Q}_{T,m} - Q_m\|_{\text{op}} \to_p 0$, we have

$$(\widehat{Q}_{T,m})^{-1} \Pi_m \psi = Q_m^{-1} \Pi_m \psi + o_p(1) \qquad \text{in } H_m.$$

Since ψ is fixed and $\bigcup_m H_m$ is dense in H, standard Galerkin arguments imply

$$Q_m^{-1} \Pi_m \psi \;\to\; Q^{-1}\psi \qquad \text{in } H.$$

Therefore,

$$\sqrt{T}\,\langle \Pi_m \widehat{\beta}_{T,m} - \beta, \psi\rangle_H = \left\langle S_T, Q^{-1}\psi \right\rangle_H + o_p(1) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t Z_t(Q^{-1}\psi) + o_p(1).$$

The sequence $\{u_t Z_t(Q^{-1}\psi)\}$ is strictly stationary, mean zero, and α-mixing with $\sum_{k\geq 1}\alpha(k)^{\delta/(2+\delta)} < \infty$. Moreover, by Lemma 1,

$$\mathbb{E}|u_t Z_t(Q^{-1}\psi)|^{2+\delta} < \infty.$$

Hence, by a central limit theorem for strongly mixing sequences (e.g. White (2001), Theorem 5.20),

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t Z_t(Q^{-1}\psi) \xrightarrow{d} \mathcal{N}\left(0, \sum_{k\in\mathbb{Z}} \mathbb{E}\big[u_t Z_t(Q^{-1}\psi)\, u_{t-k} Z_{t-k}(Q^{-1}\psi)\big]\right).$$

By Lemma 4, the asymptotic variance equals $\langle \psi, Q^{-1}\Omega Q^{-1}\psi\rangle_H$. Combining the above results yields

$$\sqrt{T}\,\langle \Pi_m \widehat{\beta}_{T,m} - \beta, \psi\rangle_H \xrightarrow{d} \mathcal{N}\big(0,\, \langle \psi, Q^{-1}\Omega Q^{-1}\psi\rangle_H\big),$$

which completes the proof of Theorem 1(ii). □

B Proof of Proposition 1

This appendix provides a proof to Proposition 1 in the main text.

Lemma 5. *Assume that $\lambda_T \to 0$:*

$$\sup_{\|f\|_H \leq B} \lambda_T \langle f, Pf\rangle_H = \lambda_T \cdot C(B) \to 0.$$

Proof. The estimation occurs in the finite-dimensional subspace $H_m \subset H$ (step functions with m bins). Its dimension is mJ. The set

$$\{f \in H_m : \|f\|_H \leq B\}$$

is compact (closed and bounded in finite dimensions). The quadratic form $f \mapsto \langle f, Pf\rangle_H$ is continuous on H_m. By compactness and continuity:

$$C(B) := \sup_{\|f\|_H \leq B} \langle f, Pf\rangle_H < \infty.$$

Since $\lambda_T \to 0$:

$$\sup_{\|f\|_H \leq B} \lambda_T \langle f, Pf\rangle_H = \lambda_T \cdot C(B) \to 0.$$

This shows the penalty term vanishes uniformly on bounded sets, ensuring the regularized

estimator behaves like the unpenalized one asymptotically. □

We now prove Proposition 1.

Proof. We build upon the Hilbert space framework established in Theorem 1. The key is to define a penalty operator P on the space H whose action is approximated by the matrix $\boldsymbol{A}_m$ on the subspace H_m. For cubic smoothing splines, the discrete penalty matrix $\mathbf{A}_m$ is a standard second-difference approximation to the integral of the squared second derivative. The approximation error $\left|\langle f, Pf\rangle_H - \theta^\top \mathbf{A}_m \theta\right|$ for $f = \Pi_m \theta$ is $O(m^{-2})$ uniformly over $\|f\|_H \leq B$ (Eilers and Marx, 1996, Section 3). This error is asymptotically negligible because $m \to \infty$ and $\lambda_T \to 0$.

Since $\hat{f}_{T,\lambda}$ minimizes the penalized sample criterion $S_{T,\lambda}(f) = S_T(f) + \lambda_T \langle f, Pf\rangle_H$, we have:

$$S_{T,\lambda}(\hat{f}_{T,\lambda}) \leq S_{T,\lambda}(\beta) = S_T(\beta) + \lambda_T \langle \beta, P\beta\rangle_H \xrightarrow{p} L(\beta).$$

Hence, for any $\epsilon > 0$, there exists T_0 such that for all $T > T_0$:

$$S_{T,\lambda}(\hat{f}_{T,\lambda}) \leq L(\beta) + \epsilon \quad \text{with high probability.}$$

Assume, for contradiction, that $\|\hat{f}_{T,\lambda}\|_H \to \infty$. Then, by the coercivity of L (since Q is positive definite):

$$L(\hat{f}_{T,\lambda}) \to \infty$$

From the uniform convergence results: - $S_T(f) \xrightarrow{p} L(f)$ uniformly on compact sets (by Theorem 1), - $\lambda_T \langle f, Pf\rangle_H \to 0$ uniformly on bounded sets (as shown earlier), it follows that:

$$S_{T,\lambda}(f) \xrightarrow{p} L(f) \quad \text{uniformly on compact sets.}$$

Now, for large T,

$$|S_{T,\lambda}(\hat{f}_{T,\lambda}) - L(\hat{f}_{T,\lambda})| < \epsilon \quad \text{and} \quad S_{T,\lambda}(\hat{f}_{T,\lambda}) \leq L(\beta) + \epsilon,$$

which implies:

$$L(\hat{f}_{T,\lambda}) < S_{T,\lambda}(\hat{f}_{T,\lambda}) + \epsilon \leq L(\beta) + 2\epsilon.$$

But this contradicts $L(\hat{f}_{T,\lambda}) \to \infty$. Therefore, the assumption that $\|\hat{f}_{T,\lambda}\|_H \to \infty$ must be

false. Thus, $\|\hat{f}_{T,\lambda}\|_H$ is bounded with probability tending to 1. That is, there exists $B > 0$ such that:

$$\hat{f}_{T,\lambda} \in \mathcal{F}_m(B) = \{f \in H_m : \|f\|_H \leq B, \langle f, Pf\rangle_H \leq B\}.$$

Since $\hat{f}_{T,\lambda}$ is bounded, it lies in the set $\{\|f\|_H \leq B\}$. On this set, we have:

$$\sup_{\|f\|_H \leq B} |S_{T,\lambda}(f) - L(f)| \leq \sup_{\|f\|_H \leq B} |S_T(f) - L(f)| + \sup_{\|f\|_H \leq B} \lambda_T \langle f, Pf\rangle_H.$$

The second term on the right-hand side converge to 0 in probability as $\lambda_T \to 0$. For the first term, the uniform convergence argument mirrors Theorem 1, but the entropy of the regularized function class $\mathcal{F}_m(B)$ is governed by the effective dimension $d_{\text{eff}}(\lambda_T)$ rather than the nominal dimension mJ.

The penalized estimator minimizes $S_T(f) + \lambda_T \theta^\top \boldsymbol{A}_m \theta$. The penalty term $\theta^\top \boldsymbol{A}_m \theta$ defines a seminorm on H_m. The set of functions satisfying $\theta^\top \boldsymbol{A}_m \theta \leq B$ is not a ball but an ellipsoid. The shape of this ellipsoid is determined by the eigenvalues of $\boldsymbol{A}_m$. The effective function class $\mathcal{F}_m(B) = \{f \in H_m : \|f\|_H \leq B, \langle f, Pf\rangle_H \leq B\}$ is the intersection of the L^2 ball with this ellipsoid. The effective degrees of freedom is defined as:

$$d_{\text{eff}}(\lambda_T) = \text{tr}(\boldsymbol{H}_\lambda) = \text{tr}\left(\boldsymbol{X}^\star (\boldsymbol{X}^{\star\top} \boldsymbol{X}^\star + \lambda_T T \boldsymbol{A}_m)^{-1} \boldsymbol{X}^{\star\top}\right).$$

Consider the spectral decomposition. The matrix $\boldsymbol{X}^{\star\top} \boldsymbol{X}^\star$ is the empirical version of the operator Q, and $\boldsymbol{A}_m$ is the penalty operator. The trace:

$$d_{\text{eff}}(\lambda_T) = \sum_{i=1}^{mJ} \frac{\mu_i}{\mu_i + \lambda_T T \nu_i}$$

where μ_i are the eigenvalues of $\boldsymbol{X}^{\star\top} \boldsymbol{X}^\star$ and ν_i are the eigenvalues of $\boldsymbol{A}_m$.

The metric entropy of the class $\mathcal{F}_m(B)$ satisfies:

$$\log N\left(\epsilon, \mathcal{F}_m(B), \|\cdot\|_H\right) \lesssim d_{\text{eff}}(\lambda_T) \cdot \log(1/\epsilon).$$

This bound follows from standard results on covering numbers of ellipsoids (Vershynin, 2018). Specifically, the penalty term $\theta^\top \mathbf{A}_m \theta$ defines a seminorm whose unit ball is an ellipsoid with effective dimension $d_{\text{eff}}(\lambda_T) = \text{tr}(\mathbf{H}_\lambda)$. Intersecting with the L^2 ball $\{\|\theta\|^2 \leq (m/\tau)B\}$ yields a set that is contained in a Euclidean ball of radius $\sqrt{d_{\text{eff}}(\lambda_T)}$ after a suitable

linear transformation (Claeskens et al., 2009). Hence the ϵ-covering number is at most $(3\sqrt{d_{\text{eff}}(\lambda_T)}/\epsilon)^{d_{\text{eff}}(\lambda_T)}$, giving the claimed logarithmic bound.

Following the same ϵ-net and union bound argument as in Theorem 1, the probability of large uniform deviations is controlled by:

$$\mathbb{P}\left(\sup_{f\in\mathcal{F}_m(B)} |S_T(f) - L(f)| > \delta\right) \lesssim N(\epsilon, \mathcal{F}_m(B), \|\cdot\|_H) \cdot \exp\left(-C\frac{T\delta^2}{G}\right).$$

Substituting the entropy bound:

$$\mathbb{P}(\cdots) \lesssim \exp\left(d_{\text{eff}}(\lambda_T)\log(1/\epsilon) - C\frac{T\delta^2}{G}\right).$$

For the probability to vanish, the exponent must go to $-\infty$. This occurs if:

$$d_{\text{eff}}(\lambda_T)\log(1/\epsilon) = o(T).$$

This is precisely the condition in Proposition 1. Under it, the uniform convergence $\sup_{f\in\mathcal{F}_m(B)} |S_T(f) - L(f)| \xrightarrow{p} 0$ holds by a standard maximal inequality for α-mixing processes (Doukhan, 1994, Theorem 3.4). The penalty term $\lambda_T\langle f, Pf\rangle_H$ vanishes uniformly on bounded sets because $\lambda_T \to 0$ and $\langle f, Pf\rangle_H$ is bounded on $\mathcal{F}_m(B)$ by construction. Hence $S_{T,\lambda}(f)$ converges uniformly to $L(f)$ on $\mathcal{F}_m(B)$.

Since $\epsilon > 0$ is fixed in the net construction, the sufficient condition is:

$$d_{\text{eff}}(\lambda_T)\log T = o(T).$$

Under condition (3), the uniform convergence holds:

$$\sup_{f\in\mathcal{F}_m(B)} |S_T(f) - L(f)| \xrightarrow{p} 0.$$

Therefore, $S_{T,\lambda}(f)$ converges uniformly to $L(f)$ on $\{\|f\|_H \le B\}$. Since $\hat{f}_{T,\lambda}$ minimizes $S_{T,\lambda}(f)$ and $L(f)$ is minimized at $f = \beta$, it follows that $\hat{f}_{T,\lambda} \to \beta$ in probability.

Multiplying both sides by $\sqrt{T}$ and rearranging gives

$$\sqrt{T}(\hat{f}_{T,\lambda} - \beta) = (\widehat{Q}_{T,m} + \lambda_T P)^{-1}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t Z_t(\cdot) - \sqrt{T}\lambda_T P\beta\right).$$

We now derive the asymptotic distribution of fixed linear functionals. Fix any $\psi \in H$.

Taking inner products with ψ yields

$$\sqrt{T}\,\langle \hat{f}_{T,\lambda}-\beta,\psi\rangle_H = \Big\langle \frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t Z_t(\cdot), (\widehat{Q}_{T,m}+\lambda_T P)^{-1}\Pi_m\psi\Big\rangle_H - \sqrt{T}\lambda_T\langle P\beta, (\widehat{Q}_{T,m}+\lambda_T P)^{-1}\Pi_m\psi\rangle_H.$$

Since $\sqrt{T}\lambda_T \to 0$ and $(\widehat{Q}_{T,m}+\lambda_T P)^{-1}$ is uniformly bounded in operator norm with probability tending to one, the second term is $o_p(1)$. Uniform boundedness of $(\widehat{Q}_{T,m}+\lambda_T P)^{-1}$ follows because $\widehat{Q}_{T,m}$ converges to Q in operator norm, and Q is coercive: $\langle f, Qf\rangle_H \geq c\|f\|_H^2$ with $c > 0$. Adding the non-negative operator $\lambda_T P$ can only increase the smallest eigenvalue. Hence, with probability approaching one, $\|(\widehat{Q}_{T,m}+\lambda_T P)^{-1}\|_{\text{op}} \leq 2/c$. The term $\sqrt{T}\lambda_T\langle P\beta,(\cdot)\rangle_H$ is then bounded by $\sqrt{T}\lambda_T\|P\beta\|_H\cdot(2/c)$, which converges to zero because $\sqrt{T}\lambda_T \to 0$.

Moreover, since $\widehat{Q}_{T,m} \to_p Q$ and $\lambda_T \to 0$, we have

$$(\widehat{Q}_{T,m}+\lambda_T P)^{-1}\Pi_m\psi = Q^{-1}\psi + o_p(1) \qquad \text{in } H.$$

Therefore,

$$\sqrt{T}\,\langle \hat{f}_{T,\lambda}-\beta,\psi\rangle_H = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t Z_t(Q^{-1}\psi) + o_p(1).$$

The sequence $\{u_t Z_t(Q^{-1}\psi)\}$ is strictly stationary, mean zero, and α-mixing with $\sum_{k\geq 1}\alpha(k)^{\delta/(2+\delta)} < \infty$. By Lemma 1,

$$\mathbb{E}|u_t Z_t(Q^{-1}\psi)|^{2+\delta} < \infty.$$

Hence, by a central limit theorem for strongly mixing sequences (e.g. White (2001), Theorem 5.20),

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t Z_t(Q^{-1}\psi) \xrightarrow{d} \mathcal{N}\big(0,\, \langle\psi, Q^{-1}\Omega Q^{-1}\psi\rangle_H\big).$$

This establishes the asymptotic normality of fixed linear functionals and completes the proof of Proposition 1.

□

C Proof of Corollary 1

Here, we present the proof of Bernstein-Von-Mise result.

Proof. Given σ^2, the posterior on H_m is Gaussian with mean $\widehat{f}_{T,\lambda}$ and covariance

$$\Sigma_{T,\lambda}(\sigma^2) = \Big(\tfrac{T}{\sigma^2} Q_{T,m} + \lambda_T A_m\Big)^{-1} = \tfrac{\sigma^2}{T}\Big(Q_{T,m} + \tfrac{\sigma^2}{T}\lambda_T A_m\Big)^{-1}.$$

We now establish a Bernstein–von Mises (BvM) result for the linear functional $L(f) = \langle f, \psi\rangle_H$. A general BvM theorem for semiparametric models under strong mixing exists in the literature. Their conditions require: (i) local asymptotic normality of the likelihood, (ii) a prior that concentrates at a suitable rate, and (iii) existence of a consistent estimator. In our setting, the sieve H_m with step functions and the Gaussian prior $\beta_j \mid \tau_j^2 \sim \mathcal{N}(0, \tau_j^2 \mathbf{A}_{m_j}^{-1})$ satisfies these conditions provided the effective degrees of freedom satisfy $d_{\text{eff}}(\lambda_T)\log T = o(T)$ and $\lambda_T \to 0$ (Castillo and Nickl, 2014).

Hence, for $L(f) = \langle f, \psi\rangle_H$,

$$Var\big(L(f) \mid \mathbf{y}, \sigma^2\big) = \langle \psi, \Sigma_{T,\lambda}(\sigma^2)\psi\rangle_H = \tfrac{\sigma^2}{T}\Big\langle \psi, \Big(Q_{T,m} + \tfrac{\sigma^2}{T}\lambda_T A_m\Big)^{-1}\psi\Big\rangle_H.$$

Let $E_T(\sigma^2) := (Q_{T,m} - Q) + (\sigma^2/T)\lambda_T A_m$. Under $Q_{T,m} \to Q$ in operator norm, $Q \succeq cI$, $\|A_m\| = O(1)$, and $\lambda_T \to 0$, we have $\|E_T(\sigma^2)\| \to 0$ uniformly over stochastically bounded σ^2. By the resolvent identity,

$$\Big(Q_{T,m} + \tfrac{\sigma^2}{T}\lambda_T A_m\Big)^{-1} - Q^{-1} = -\,Q^{-1}E_T(\sigma^2)\Big(Q + E_T(\sigma^2)\Big)^{-1} = o_P(1).$$

Because of conjugacy, the conditional posterior distribution of σ^2 is

$$\sigma^2 \mid \mathbf{y} \;\sim\; \text{IG}\left(a_0 + \tfrac{T}{2},\; b_0 + \tfrac{1}{2}\sum_{t=1}^{T}(y_t - Z_t(\widehat{f}_{T,\lambda}))^2\right).$$

Let

$$S_T = \frac{1}{T}\sum_{t=1}^{T}(y_t - Z_t(\widehat{f}_{T,\lambda}))^2.$$

By a law of large numbers and consistency of $\widehat{f}_{T,\lambda}$, we have $S_T \xrightarrow{P} \sigma_0^2$. Hence the scale parameter in the inverse–Gamma posterior is asymptotically $b_0 + T\sigma_0^2/2$, while the shape parameter is $a_0 + T/2$. For large T, the inverse–Gamma distribution is sharply concentrated, and one can show by a delta–method expansion of its mean or mode that

$$\mathbb{E}[\sigma^2 \mid \mathbf{y}] = \sigma_0^2 + O_P(T^{-1}), \qquad Var(\sigma^2 \mid \mathbf{y}) = \frac{2\sigma_0^4}{T} + o(T^{-1}).$$

Therefore, after centering and $\sqrt{T}$–rescaling,

$$\sqrt{T}(\sigma^2 - \sigma_0^2)|\mathbf{y} \;\Rightarrow\; \mathcal{N}(0,\, 2\sigma_0^4),$$

which implies

$$\sigma^2 = \sigma_0^2 + O_P(T^{-1/2})$$

under the posterior. Consequently, replacing σ^2 by σ_0^2 in the covariance expression perturbs the variance of $L(f)$ only by $o_P(T^{-1})$. Combining this with the variance and dominated convergence yields the *unconditional* posterior variance

$$Var\big(L(f) \mid \mathbf{y}\big) = \tfrac{\sigma_0^2}{T}\,\langle \psi, Q^{-1}\psi\rangle_H + o_P(T^{-1}) = \tfrac{\sigma^2}{T}\,\langle \psi, Q^{-1}\psi\rangle_H + o_P(T^{-1}),$$

which is asymptotically deterministic.

By a standard result on mixtures of normal distributions with a mixing distribution that converges to a point (e.g., van der Vaart (1998)), the marginal posterior of $\sqrt{T}(L(f) - L(\beta))$ inherits asymptotic normality.

The posterior distribution of the scalar functional can be written as

$$\sqrt{T}\{L(f) - L(\beta)\} = \sqrt{T}\{L(f) - L(\hat{f}_{T,\lambda})\} + \sqrt{T}\{L(\hat{f}_{T,\lambda}) - L(\beta)\}.$$

Conditional on the data, $\hat{f}_{T,\lambda}$ is fixed, so the posterior fluctuation $L(f) - L(\hat{f}_{T,\lambda})$ is independent of $\hat{f}_{T,\lambda}$ in the posterior distribution.

Conditional on σ^2, the first term is Gaussian with mean zero and variance $Var(L(f) \mid \mathbf{y}, \sigma^2)$. By the previous step, this variance converges in probability to $\sigma_0^2\langle \psi, Q^{-1}\psi\rangle_H/T$. The second term converges in distribution to $\mathcal{N}(0, \langle \psi, Q^{-1}\Omega Q^{-1}\psi\rangle_H)$ by Proposition 1 (which verified the required mixing CLT for the score process) Because the two terms are independent in the posterior, standard Gaussian shift arguments (see Bickel and Kleijn (2012)) imply that the posterior of $\sqrt{T}(L(f) - L(\beta))$ converges to a Gaussian distribution with the same variance as the sampling distribution of $\sqrt{T}(L(\hat{f}_{T,\lambda}) - L(\beta))$. This is the Bernstein–von Mises result stated in Corollary 1. □

The result is straightforward in our setting. On the sieve space H_m, the likelihood is Gaussian and the spline prior is Gaussian, so the posterior for any fixed linear functional is also Gaussian. Its center is the penalized estimator, and its variance is of order T^{-1}. Proposition 1 gives the asymptotic distribution of the center, while the condition $\sqrt{T}\lambda_T \to 0$ ensures that the prior bias is asymptotically negligible. It then follows that the Bernstein–von Mises

result is a direct first-order consequence of the main asymptotic theory.

D Additional Monte Carlo Results

This appendix reports additional Monte Carlo results for Case 1 (univariate, $J = 1$) and for smaller sample sizes. Tables A.1 and A.2 report results for $T \in \{100, 200\}$ under the same frequency-mismatch regimes as in the main text. We compare four estimators: (i) unregularized U-MIDAS (OLS), (ii) adaptive group LASSO, (iii) cubic smoothing-spline regularization with a fixed smoothing parameter $\lambda_T = T^{-3/4}$ (chosen to satisfy the conditions in Proposition 1), and (iv) cubic smoothing-spline regularization with an adaptive smoothing parameter implemented through the hierarchical prior on τ_j^2, so that $\lambda_j = \sigma^2/(T\tau_j^2)$ varies across MCMC iterations.

For each design we report the integrated squared error (ISE) of the implied step-function estimator, the average Monte Carlo variance of the estimated coefficients across bins, and the mean condition number of $\boldsymbol{X}^{\star\top}\boldsymbol{X}^{\star}$ across replications. The results mirror the patterns in the main text: unregularized U-MIDAS performs well when $m = o(T/\log T)$ but deteriorates under extreme mismatch, while cubic smoothing-spline regularization remains stable and substantially reduces estimation error and numerical instability. Adaptive group LASSO typically improves upon no shrinkage, but its performance worsens as m becomes large relative to T.

In addition, Tables A.4 and A.5 report the empirical coverage of 95% highest posterior density credible intervals for linear functionals under Case 2 and cubic smoothing spline with $T = 200$ and $T = 100$ respectively. Again, we consider the linear functionals based on early-lag, late-lag, and oscillatory weighting functions specified in Footnote 1 in the main text. Results show that the empirical coverage rate of 95% credible intervals are fairly close to the nominal coverage rate under different degrees of frequency mismatch. It is clear that cubic smoothing spline is able to recover stable coverage of credible intervals even under extreme frequency mismatch.

Table A.1. Results of Monte Carlo Experiments under Case 1 (Univariate) with $T = 100$

$m =$	$2\sqrt{T}$		$5\sqrt{T}$		$T^{0.75}$		$2T$	
$\rho =$	0	0.5	0	0.5	0	0.5	0	0.5
True lag-weight function: Smooth								
No shrinkage								
ISE	0.2556	0.3327	1.0464	1.3626	0.4854	0.6274	6.1910	6.4709
Mean variance of $\hat{\boldsymbol{\beta}}$	0.2555	0.3325	1.0461	1.3630	0.4854	0.6276	5.6257	5.9030
Adaptive group LASSO								
ISE	0.2301	0.2928	0.6292	0.7407	0.3885	0.4740	1.5884	1.6536
Mean variance of $\hat{\boldsymbol{\beta}}$	0.2102	0.2599	0.4522	0.5031	0.3234	0.3772	0.6291	0.6243
Cubic smoothing spline (fixed λ_T)								
ISE	0.0708	0.0917	0.1355	0.1797	0.0966	0.1244	0.4241	0.5492
Mean variance of $\hat{\boldsymbol{\beta}}$	0.0686	0.0894	0.1355	0.1797	0.0962	0.1239	0.4243	0.5491
Cubic smoothing spline								
ISE	0.0532	0.0649	0.0555	0.0698	0.0529	0.0629	0.1072	0.1321
Mean variance of $\hat{\boldsymbol{\beta}}$	0.0391	0.0477	0.0492	0.0625	0.0431	0.0515	0.1069	0.1317
True lag-weight function: Localized								
No shrinkage								
ISE	0.2556	0.3327	1.0464	1.3626	0.4854	0.6274	5.8038	6.0837
Mean variance of $\hat{\boldsymbol{\beta}}$	0.2555	0.3325	1.0461	1.3630	0.4854	0.6276	5.4403	5.7188
Adaptive group LASSO								
ISE	0.2171	0.2704	0.5136	0.5898	0.3462	0.4135	1.1017	1.1454
Mean variance of $\hat{\boldsymbol{\beta}}$	0.1819	0.2183	0.3311	0.3528	0.2604	0.2930	0.4114	0.4040
Cubic smoothing spline (fixed λ_T)								
ISE	0.1724	0.1934	0.1421	0.1864	0.1297	0.1567	0.4241	0.5491
Mean variance of $\hat{\boldsymbol{\beta}}$	0.0739	0.0950	0.1360	0.1804	0.0993	0.1266	0.4243	0.5491
Cubic smoothing spline								
ISE	0.1631	0.2113	0.1518	0.1941	0.1591	0.2018	0.1319	0.1626
Mean variance of $\hat{\boldsymbol{\beta}}$	0.1006	0.1208	0.0967	0.1187	0.0984	0.1156	0.1196	0.1467
Mean $\kappa(\boldsymbol{X}^{\star\top}\boldsymbol{X}^{\star})$	60	60	535	535	156	156	3.0×10^{20}	3.0×10^{20}

Table A.2. Results of Monte Carlo Experiments under Case 1 (Univariate) with $T = 200$

$m =$	$2\sqrt{T}$		$5\sqrt{T}$		$T^{0.75}$		$2T$	
$\rho =$	0	0.5	0	0.5	0	0.5	0	0.5
True lag-weight function: Smooth								
No shrinkage								
ISE	0.1647	0.2165	0.5635	0.7448	0.3654	0.4903	5.9130	6.2309
Mean variance of $\hat{\boldsymbol{\beta}}$	0.1647	0.2163	0.5632	0.7446	0.3653	0.4902	5.3758	5.6938
Adaptive group LASSO								
ISE	0.1555	0.1999	0.4179	0.5159	0.3057	0.3906	1.5465	1.6109
Mean variance of $\hat{\boldsymbol{\beta}}$	0.1451	0.1836	0.3418	0.3999	0.2657	0.3264	0.5563	0.5395
Cubic smoothing spline (fixed λ_T)								
ISE	0.0513	0.0678	0.0989	0.1312	0.0803	0.1076	0.4060	0.5397
Mean variance of $\hat{\boldsymbol{\beta}}$	0.0509	0.0673	0.0987	0.1310	0.0803	0.1076	0.4061	0.5398
Cubic smoothing spline								
ISE	0.0325	0.0391	0.0357	0.0441	0.0322	0.0412	0.0936	0.1168
Mean variance of $\hat{\boldsymbol{\beta}}$	0.0247	0.0299	0.0335	0.0414	0.0294	0.0377	0.0936	0.1168
True lag-weight function: Localized								
No shrinkage								
ISE	0.1647	0.2165	0.5635	0.7448	0.3654	0.4903	5.5454	5.8633
Mean variance of $\hat{\boldsymbol{\beta}}$	0.1647	0.2163	0.5632	0.7446	0.3653	0.4902	5.1874	5.5028
Adaptive group LASSO								
ISE	0.1471	0.1869	0.3686	0.4426	0.2827	0.3515	1.0881	1.1327
Mean variance of $\hat{\boldsymbol{\beta}}$	0.1336	0.1652	0.2763	0.3123	0.2263	0.2683	0.3672	0.3583
Cubic smoothing spline (fixed λ_T)								
ISE	0.0749	0.0906	0.0994	0.1315	0.0829	0.1100	0.4060	0.5397
Mean variance of $\hat{\boldsymbol{\beta}}$	0.0520	0.0684	0.0988	0.1311	0.0805	0.1076	0.4061	0.5398
Cubic smoothing spline								
ISE	0.0777	0.0994	0.0723	0.0912	0.0753	0.0971	0.0966	0.1207
Mean variance of $\hat{\boldsymbol{\beta}}$	0.0592	0.0729	0.0562	0.0683	0.0568	0.0705	0.0958	0.1196
Mean $\kappa(\boldsymbol{X}^{\star\top}\boldsymbol{X}^{\star})$	68	68	411	411	213	213	3.6×10^{20}	3.6×10^{20}

Table A.3. Results of Monte Carlo Experiments under Case 2 (Multivariate) with $T = 100$

$m =$	$2\sqrt{T}$		$5\sqrt{T}$		$T^{0.75}$		$2T$	
$\rho =$	0	0.5	0	0.5	0	0.5	0	0.5
True lag-weight function: Smooth								
No shrinkage								
ISE	1.692	2.216	900.944	901.522	44.308	54.756	571.107	571.160
Mean variance of $\hat{\boldsymbol{\beta}}$	0.564	0.739	300.154	300.347	14.763	18.253	189.101	189.119
Adaptive group LASSO								
ISE	1.044	1.254	3.135	3.394	1.972	2.193	5.237	5.292
Mean variance of $\hat{\boldsymbol{\beta}}$	0.330	0.391	0.848	0.916	0.586	0.641	0.457	0.473
Cubic smoothing spline (fixed λ_T)								
ISE	0.243	0.307	0.503	0.655	0.342	0.434	1.793	2.331
Mean variance of $\hat{\boldsymbol{\beta}}$	0.079	0.101	0.168	0.219	0.114	0.144	0.597	0.777
Cubic smoothing spline								
ISE	0.177	0.209	0.189	0.229	0.172	0.203	0.418	0.509
Mean variance of $\hat{\boldsymbol{\beta}}$	0.049	0.058	0.059	0.072	0.051	0.060	0.139	0.169
True lag-weight function: Localized								
No shrinkage								
ISE	1.692	2.216	900.352	900.930	44.301	54.757	569.330	569.383
Mean variance of $\hat{\boldsymbol{\beta}}$	0.564	0.739	300.001	300.194	14.761	18.253	188.958	188.976
Adaptive group LASSO								
ISE	0.871	1.031	2.170	2.323	1.465	1.619	3.464	3.515
Mean variance of $\hat{\boldsymbol{\beta}}$	0.265	0.307	0.575	0.606	0.418	0.451	0.314	0.329
Cubic smoothing spline (fixed λ_T)								
ISE	0.542	0.606	0.537	0.688	0.450	0.542	1.793	2.331
Mean variance of $\hat{\boldsymbol{\beta}}$	0.107	0.128	0.172	0.223	0.127	0.158	0.597	0.777
Cubic smoothing spline								
ISE	0.586	0.731	0.531	0.651	0.562	0.682	0.513	0.623
Mean variance of $\hat{\boldsymbol{\beta}}$	0.156	0.186	0.141	0.168	0.150	0.175	0.163	0.197
Mean $\kappa(\boldsymbol{X}^{\star\top}\boldsymbol{X}^{\star})$	408	408	4.0×10^{19}	4.0×10^{19}	98202	98202	4.2×10^{20}	4.2×10^{20}

Table A.4. Empirical coverage of 95% credible intervals for linear functionals under cubic smoothing spline (Case 2, $T = 200$, $\rho = 0.5$)

		Early-lag			Late-lag			Oscillatory		
True function		x_1	x_2	x_3	x_1	x_2	x_3	x_1	x_2	x_3
$m = 2\sqrt{T}$										
Smooth	Coverage	0.942	0.952	0.934	0.948	0.958	0.950	0.968	0.968	0.946
	Avg. length	4.997	5.163	4.790	5.327	5.489	5.097	6.245	6.411	5.962
	MC std of $\hat{L}(\beta)$	1.255	1.316	1.287	1.369	1.389	1.350	1.471	1.502	1.567
	Avg. $d_{\text{eff}}(\lambda_T)$	17.701								
Local	Coverage	0.952	0.964	0.932	0.938	0.958	0.942	0.960	0.970	0.942
	Avg. length	5.347	5.529	5.130	6.110	6.293	5.852	7.150	7.365	6.833
	MC std of $\hat{L}(\beta)$	1.350	1.411	1.365	1.582	1.576	1.533	1.739	1.767	1.789
	Avg. $d_{\text{eff}}(\lambda_T)$	39.086								
$m = 5\sqrt{T}$										
Smooth	Coverage	0.954	0.958	0.952	0.942	0.938	0.952	0.940	0.940	0.952
	Avg. length	12.535	12.847	11.921	14.284	14.630	13.579	16.367	16.808	15.629
	MC std of $\hat{L}(\beta)$	3.215	3.133	2.978	3.669	3.735	3.288	4.179	4.292	3.905
	Avg. $d_{\text{eff}}(\lambda_T)$	24.173								
Local	Coverage	0.946	0.944	0.954	0.944	0.946	0.962	0.938	0.940	0.954
	Avg. length	13.231	13.572	12.579	15.403	15.808	14.645	17.684	18.245	16.920
	MC std of $\hat{L}(\beta)$	3.456	3.381	3.083	3.998	4.096	3.488	4.557	4.784	4.254
	Avg. $d_{\text{eff}}(\lambda_T)$	38.082								
$m = T^{0.75}$										
Smooth	Coverage	0.966	0.962	0.956	0.966	0.964	0.946	0.966	0.960	0.958
	Avg. length	9.403	9.649	8.981	10.496	10.816	10.031	12.150	12.385	11.567
	MC std of $\hat{L}(\beta)$	2.275	2.304	2.268	2.416	2.589	2.534	2.912	3.012	2.852
	Avg. $d_{\text{eff}}(\lambda_T)$	21.095								
Local	Coverage	0.968	0.974	0.944	0.956	0.970	0.950	0.962	0.960	0.954
	Avg. length	10.009	10.269	9.560	11.530	11.921	11.039	13.412	13.743	12.777
	MC std of $\hat{L}(\beta)$	2.471	2.393	2.497	2.768	2.859	2.819	3.193	3.322	3.147
	Avg. $d_{\text{eff}}(\lambda_T)$	37.994								
$m = 2T$										
Smooth	Coverage	0.938	0.940	0.944	0.964	0.944	0.934	0.950	0.956	0.952
	Avg. length	75.523	78.120	72.336	91.625	94.582	87.675	105.254	108.098	100.732
	MC std of $\hat{L}(\beta)$	19.204	20.641	18.777	22.409	25.056	22.192	25.504	27.202	25.812
	Avg. $d_{\text{eff}}(\lambda_T)$	61.294								
Local	Coverage	0.938	0.944	0.946	0.964	0.946	0.938	0.952	0.952	0.954
	Avg. length	75.835	78.446	72.648	92.045	95.030	88.091	105.733	108.611	101.195
	MC std of $\hat{L}(\beta)$	19.307	20.738	18.840	22.619	25.143	22.310	25.641	27.273	25.884
	Avg. $d_{\text{eff}}(\lambda_T)$	62.529								

Notes: Results based on 500 replications. "Coverage" is the fraction of replications where the true value lies inside the 95% highest posterior density credible interval. "Avg. length" is the average interval length. "MC std" is the Monte Carlo standard deviation of the point estimate $\hat{L}(\beta)$ across replications. "Avg. $d_{\text{eff}}(\lambda_T)$" is the average effective degrees of freedom. x_1, x_2, and x_3 respectively denote the first, second, and third covariates in Case 2.

Table A.5. Empirical coverage of 95% credible intervals for linear functionals under cubic smoothing spline (Case 2, $T = 100$, $\rho = 0.5$)

		Early-lag			Late-lag			Oscillatory		
True function		x_1	x_2	x_3	x_1	x_2	x_3	x_1	x_2	x_3
$m = 2\sqrt{T}$										
Smooth	Coverage	0.950	0.938	0.964	0.964	0.964	0.968	0.956	0.958	0.948
	Avg. length	5.339	5.493	5.089	5.381	5.460	5.140	6.471	6.651	6.152
	MC std of $\hat{L}(\beta)$	1.311	1.433	1.244	1.249	1.297	1.230	1.583	1.645	1.546
	Avg. $d_{\text{eff}}(\lambda_T)$	13.868								
Local	Coverage	0.968	0.952	0.962	0.956	0.954	0.968	0.966	0.956	0.954
	Avg. length	6.321	6.482	5.991	6.845	6.999	6.613	8.071	8.411	7.726
	MC std of $\hat{L}(\beta)$	1.467	1.649	1.432	1.641	1.673	1.559	2.015	2.073	1.845
	Avg. $d_{\text{eff}}(\lambda_T)$	27.729								
$m = 5\sqrt{T}$										
Smooth	Coverage	0.934	0.954	0.952	0.946	0.956	0.966	0.940	0.960	0.974
	Avg. length	13.065	13.413	12.440	14.387	14.750	13.693	16.528	16.858	15.759
	MC std of $\hat{L}(\beta)$	3.401	3.384	3.099	3.578	3.757	3.197	4.360	4.243	3.557
	Avg. $d_{\text{eff}}(\lambda_T)$	17.589								
Local	Coverage	0.932	0.946	0.948	0.960	0.946	0.958	0.946	0.948	0.964
	Avg. length	15.022	15.389	14.277	16.979	17.541	16.208	19.605	20.145	18.821
	MC std of $\hat{L}(\beta)$	3.889	3.962	3.549	4.177	4.445	3.902	5.205	5.197	4.372
	Avg. $d_{\text{eff}}(\lambda_T)$	28.039								
$m = T^{0.75}$										
Smooth	Coverage	0.956	0.950	0.948	0.958	0.970	0.956	0.970	0.946	0.950
	Avg. length	8.280	8.569	7.962	8.788	9.035	8.423	10.271	10.453	9.768
	MC std of $\hat{L}(\beta)$	2.004	2.127	2.055	2.178	2.079	2.033	2.406	2.624	2.395
	Avg. $d_{\text{eff}}(\lambda_T)$	15.137								
Local	Coverage	0.968	0.956	0.946	0.950	0.976	0.962	0.960	0.956	0.938
	Avg. length	9.674	10.032	9.328	10.757	11.170	10.376	12.628	12.967	12.119
	MC std of $\hat{L}(\beta)$	2.302	2.474	2.427	2.601	2.605	2.510	3.049	3.345	3.184
	Avg. $d_{\text{eff}}(\lambda_T)$	28.274								
$m = 2T$										
Smooth	Coverage	0.940	0.954	0.936	0.940	0.932	0.956	0.938	0.946	0.938
	Avg. length	55.007	56.453	52.776	65.116	66.934	62.445	74.857	76.929	71.740
	MC std of $\hat{L}(\beta)$	14.508	14.709	13.868	16.619	17.641	15.546	19.471	20.244	18.987
	Avg. $d_{\text{eff}}(\lambda_T)$	35.542								
Local	Coverage	0.944	0.954	0.940	0.952	0.942	0.952	0.936	0.948	0.936
	Avg. length	57.082	58.602	54.819	67.766	69.681	65.002	77.953	80.106	74.684
	MC std of $\hat{L}(\beta)$	14.939	15.195	14.513	17.202	18.255	16.382	19.726	20.689	19.983
	Avg. $d_{\text{eff}}(\lambda_T)$	38.317								

Notes: See Table A.4.

E Additional SPF Uncertainty Results

This appendix reports additional results for the SPF-based macroeconomic uncertainty application that are omitted from the main text for brevity. In particular, we present results for uncertainty measured by SPF disagreement in CPI inflation and the unemployment rate, as well as detailed posterior summaries of linear functionals for all target variables.

Figures 6 and 7 report posterior inference for the linear functionals $L_{\text{tot}}(\beta)$, $L_{\text{late}}(\beta)$, and $L_{\text{tilt}}(\beta)$, corresponding to cumulative, late-quarter, and smoothly weighted timing effects. These objects are exactly the linear functionals covered by the asymptotic theory in Section 2.

Across both CPI inflation and unemployment uncertainty, posterior distributions are generally close to Gaussian, and the associated 95% highest posterior density credible intervals exhibit stable coverage. As in the real GDP case, late-quarter functionals tend to be estimated less precisely, reflecting the shorter effective time window near the end of the quarter. Finite-sample deviations from Gaussianity are occasionally visible for cumulative effects, consistent with the theoretical discussion following Corollary 1.

Table A.6 reports posterior means and 95% highest posterior density credible intervals for all linear functionals and all target variables. The table provides a comprehensive numerical summary underlying the graphical results presented in the main text and in this appendix. While individual estimates vary across predictors and target variables, the qualitative conclusions are robust: late-quarter information often plays a larger role in shaping macroeconomic uncertainty, albeit with greater posterior uncertainty, and smoothing-spline regularization yields stable and interpretable inference across specifications.

Overall, the additional results in this appendix reinforce the main findings of the paper. Regularized U-MIDAS delivers coherent inference on economically meaningful linear functionals of the lag-weight function, and the role of intra-quarter financial information in explaining macroeconomic uncertainty is broadly similar across real activity, inflation, and labor market expectations.

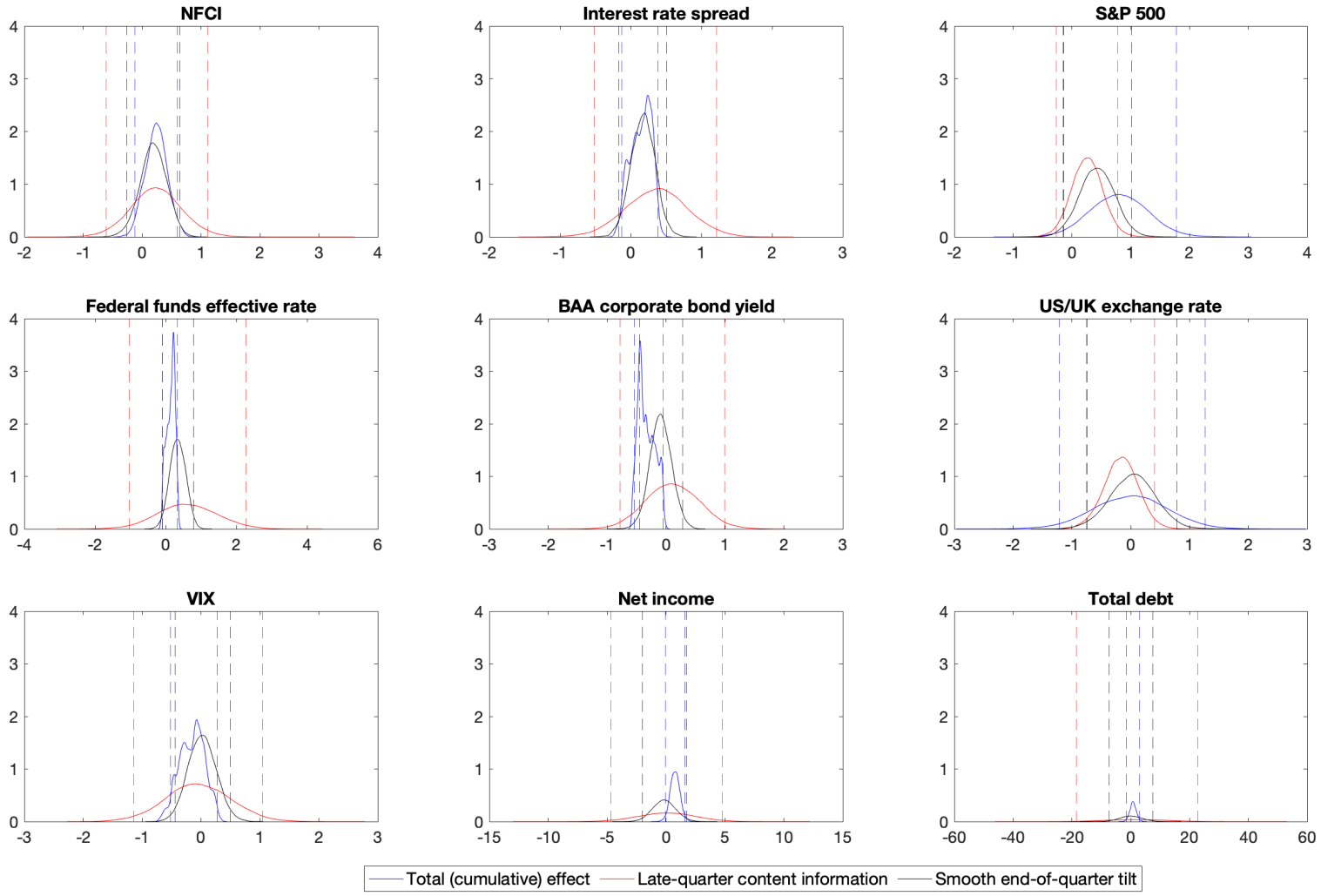

Figure 6. Uncertainty on CPI inflation: Posterior distributions of $L_{\text{tot}}(\beta)$, $L_{\text{late}}(\beta)$, and $L_{\text{tilt}}(\beta)$

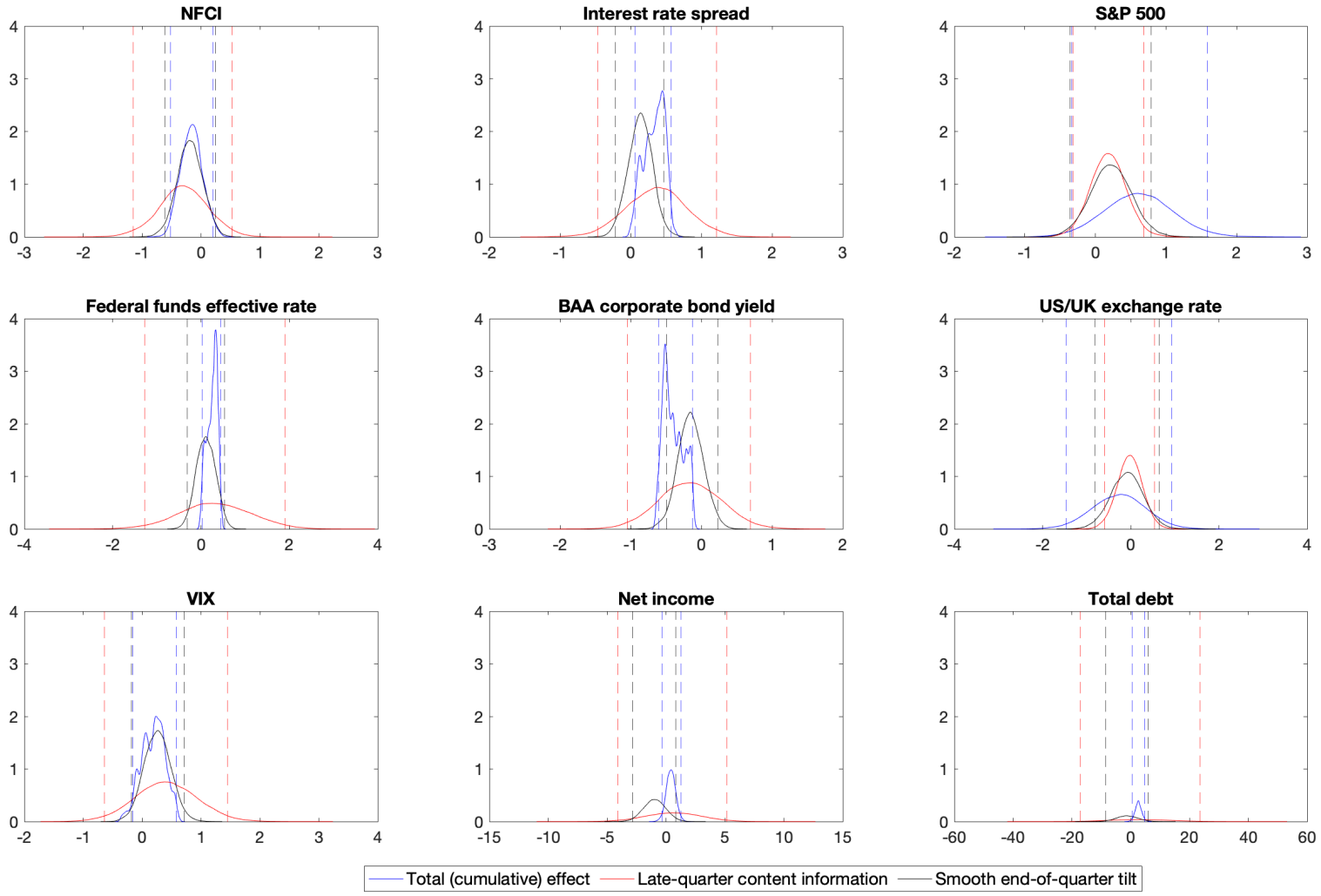

Figure 7. Uncertainty on the unemployment rate: Posterior distributions of $L_{\text{tot}}(\beta)$, $L_{\text{late}}(\beta)$, and $L_{\text{tilt}}(\beta)$

Table A.6. Posterior means and credible intervals of $L_{\text{tot}}(\beta)$, $L_{\text{late}}(\beta)$, and $L_{\text{tilt}}(\beta)$

	$L_{\text{tot}}(\beta)$			$L_{\text{late}}(\beta)$			$L_{\text{tilt}}(\beta)$		
	PM	LB	UB	PM	LB	UB	PM	LB	UB
Uncertainty on real GDP growth									
NFCI	-0.102	-0.473	0.280	-0.071	-0.913	0.822	-0.059	-0.503	0.396
Interest rate spread	0.326	0.012	0.559	0.203	-0.674	1.080	0.004	-0.347	0.353
S&P 500	0.718	-0.287	1.706	0.099	-0.431	0.617	0.237	-0.361	0.830
Federal funds effective rate	0.255	0.002	0.451	0.134	-1.500	1.798	0.039	-0.406	0.480
BAA corporate bond yield	-0.388	-0.597	-0.090	-0.038	-0.942	0.889	-0.040	-0.416	0.338
US/UK exchange rate	0.198	-1.057	1.415	0.152	-0.445	0.740	0.155	-0.615	0.899
VIX	0.183	-0.207	0.604	0.129	-0.960	1.186	0.143	-0.324	0.622
Net income	0.323	-0.480	1.195	0.122	-4.788	4.834	-1.253	-3.151	0.685
Total debt	2.138	-0.193	4.421	3.264	-18.318	23.887	-1.059	-8.583	6.533
Uncertainty on CPI inflation									
NFCI	0.240	-0.126	0.604	0.240	-0.612	1.118	0.191	-0.254	0.637
Interest rate spread	0.145	-0.130	0.384	0.364	-0.517	1.217	0.175	-0.167	0.509
S&P 500	0.803	-0.142	1.778	0.257	-0.266	0.780	0.429	-0.153	1.018
Federal funds effective rate	0.145	-0.082	0.339	0.609	-1.035	2.270	0.343	-0.088	0.786
BAA corporate bond yield	-0.330	-0.541	-0.050	0.110	-0.784	0.998	-0.089	-0.453	0.284
US/UK exchange rate	-0.007	-1.220	1.261	-0.166	-0.761	0.408	0.035	-0.741	0.786
VIX	-0.151	-0.513	0.276	-0.049	-1.139	1.042	0.021	-0.441	0.493
Net income	0.768	-0.050	1.587	0.041	-4.687	4.804	-0.164	-2.018	1.780
Total debt	0.775	-1.454	2.939	1.790	-18.376	22.765	0.047	-7.347	7.547
Uncertainty on the unemployment rate									
NFCI	-0.161	-0.519	0.203	-0.297	-1.156	0.528	-0.189	-0.612	0.243
Interest rate spread	0.333	0.063	0.570	0.362	-0.468	1.217	0.136	-0.212	0.467
S&P 500	0.615	-0.331	1.584	0.181	-0.318	0.683	0.221	-0.356	0.791
Federal funds effective rate	0.252	0.020	0.436	0.318	-1.274	1.897	0.110	-0.308	0.537
BAA corporate bond yield	-0.398	-0.607	-0.127	-0.159	-1.046	0.697	-0.150	-0.495	0.234
US/UK exchange rate	-0.267	-1.466	0.929	-0.022	-0.593	0.544	-0.084	-0.817	0.650
VIX	0.187	-0.164	0.581	0.408	-0.640	1.450	0.260	-0.184	0.716
Net income	0.417	-0.364	1.248	0.625	-4.082	5.187	-0.968	-2.823	0.835
Total debt	2.588	0.427	4.835	3.032	-17.031	23.454	-1.219	-8.494	5.978

Note: "PM" denotes the posterior mean, and "LB" and "UB" represent the lower and upper bounds of 95% highest posterior density credible interval respectively.