An Extended Score-Driven Dynamic Factor Model: Recovering Composite Indicators from the Pandemic*

Mariia Artemova, Dick van Dijk, and Evgenii Vladimirov

Erasmus University Rotterdam and Tinbergen Institute

This draft: December 14, 2025
First draft: June 16, 2025

Abstract

We propose an extended score-driven (ESD) dynamic factor model (DFM) that accommodates non-Gaussian innovations, nonlinear factor dynamics, and time-varying volatility. The main novelty of our model is a factor equation that includes both lagged and contemporaneous scores, implying that factors are not predetermined. We show that this novel model nests both the classic (parameter-driven) state-space DFM as well as the more recent score-driven DFM, bridging the gap between these two model classes. Empirically, our ESD-DFM proves useful for working with COVID-19–era observations, which have posed substantial challenges for macroeconomic modeling. For instance, while the Federal Reserve Bank of Philadelphia suspended publication of its leading index due to pandemic-related anomalies, our model remains robust to such extreme observations and enables reliable computation of the index. We further apply the ESD-DFM to The Conference Board's (TCB's) Coincident and Leading Economic Indices (CEI and LEI). When indices are constructed from the estimated factors, the unprecedented divergence between TCB's CEI and LEI observed during the post-pandemic period disappears: although the reconstructed LEI declines in 2022, it resumes an upward trajectory from late 2023 through mid-2025.

Keywords: state-space model; heavy-tailed distributions; robust filtering; observation-driven model; coincident index; leading index.

*We thank conference participants at the annual TSE Financial Econometrics Conference in 2025 for useful comments and suggestions.

E-mail addresses: artemova@ese.eur.nl (M. Artemova), djvandijk@ese.eur.nl (D. van Dijk), and vladimirov@ese.eur.nl (E. Vladimirov).

1 Introduction

The COVID-19 pandemic has posed significant challenges for macroeconomic time series modeling and forecasting. In particular, many macroeconomic variables exhibited elevated uncertainty and increased volatility during this period, straining the reliability of state-of-the-art econometric models. As a result, many institutions had to adapt their established methodologies for construction of real-time business cycle indicators or even suspend their production until further notice. In this paper, we propose a novel modeling framework that nests models widely used by policymakers prior to the pandemic, while incorporating features that have become essential in its aftermath – all within a frequentist paradigm.

The urgency for adapting macroeconomic time series models and related methods intensified after the pandemic. For example, the Federal Reserve Bank of Philadelphia, in producing its state coincident indices, treated all 2020 observations as missing data in order to *"exclude the impacts of the extreme, idiosyncratic shock from the pandemic and to preserve states' historic business cycle characteristic"*[1]. Moreover, due to the extreme impact on initial unemployment claims, the Philadelphia Fed discontinued the release of their state leading indices[2]. Similarly, the New York Fed suspended the release of its GDP nowcast for two years between September 2021 and September 2023. To adapt its model to the pandemic data, the New York Fed incorporated stochastic volatility and outlier adjustment in the latent variable dynamics (Almuzara et al., 2023). Only to accommodate these changes, the estimation is now conducted using a Bayesian approach.

These institutional responses highlight the need for modeling frameworks that can accommodate departures from Gaussianity and handle sudden increases in volatility. In this paper, we address these challenges within a frequentist paradigm for small-scale dynamic factor models (DFMs). Our approach incorporates non-Gaussian distributions and time-varying volatility, while retaining the possibility to use maximum likelihood estimation; thus offering a flexible yet tractable alternative for macroeconomic analysis in periods of heightened uncertainty.

Dynamic factor models have become workhorse models in a wide variety of macroeconomic and financial applications. For instance, many policy institutions employ a small-scale DFM based on the Kalman filter for constructing coincident and leading economic indices, tracking business cycle conditions (see the Aruoba-Diebold-Scotti index of the Philadelphia Fed, for example), nowcasting GDP (New York Fed, ECB, DNB), forecasting inflation (BoE, ECB), and financial market monitoring (e.g., CISS index by ECB), among

[1] `https://www.philadelphiafed.org/-/media/FRBP/Assets/Surveys-And-Data/coincident/Methodology.pdf`

[2] `https://www.philadelphiafed.org/surveys-and-data/regional-economic-analysis/state-leading-indexes`

many others. These, often real-time indices play important roles in shaping policy decisions and are closely watched by financial market participants.

Prior to the pandemic, the Gaussian parameter-driven DFM dominated the construction of composite indicators; see, for example, Stock and Watson (1989), Crone and Clayton-Matthews (2005), and Aruoba et al. (2009) and the examples involving central banks and other policy institutions mentioned above. The Gaussianity assumption seemed to be quite reasonable and was very convenient as it enabled maximum likelihood estimation and the application of the Kalman filter. More recent research, however, underscores the relevance of heavy-tailed innovations in macroeconomic and financial applications, particularly in the post-pandemic period (Antolín-Díaz et al., 2024; Carriero et al., 2024; Artemova, 2025). In this context, score-driven models have proven especially useful as they remain within the frequentist paradigm, are straightforward to estimate, and naturally allow for robustness. In general, they also encompass many widely used observation-driven model specifications (Creal et al., 2013; Artemova et al., 2022), deliver strong out-of-sample performance even under misspecification (Koopman et al., 2016), and, importantly, in many cases nest parameter-driven counterparts. In the univariate setting, for example, score-driven location models are observationally equivalent to ARMA processes (Artemova et al., 2022; Blasques et al., 2024), thereby nesting parameter-driven signal-plus-noise models.

In the multivariate (dynamic) factor setting, however, this equivalence often breaks down. As we illustrate in Section 2, there remains a gap between score-driven DFMs and traditional parameter-driven DFMs. Zou et al. (2025) also document this gap in the context of observed factors and time-varying loadings. Thus, while observation-driven methods offer flexibility and robustness, they do not automatically nest the parameter-driven DFMs that policymakers traditionally relied on, missing important dynamic features of the data. To resolve this issue, we introduce an extended score-driven (ESD) DFM that naturally extends the score-driven DFM considered in the literature and is such that it also nests the Gaussian parameter-driven DFM. Thus, the proposed framework bridges the gap between the two classes of models, unifying the two approaches and providing a framework that is both robust and grounded in established practice.

The proposed ESD-DFM is in the spirit of the real-time GARCH model of Smetanina (2017), that bridges stochastic volatility (SV) and GARCH models. The latter is a prominent example of an observation-driven model and is often contrasted with the parameter-driven SV model, in which volatility dynamics are driven by exogenous innovations. Similar to the real-time GARCH model, our factor updating equation incorporates the scores both at times t and $t+1$. In contrast to score-driven factor models, this implies that the factors at time $t+1$ are not predetermined at time t. However, they can be recovered once the observation at time $t+1$ becomes available. As a result, the filtering procedure is

naturally split into two steps: prediction and updating steps – analogously to the Kalman (1960) filter and the implicit score-driven filter recently proposed by (Lange et al., 2024). Finally, being observation-driven, our model allows to easily accommodate non-Gaussian distributions and time-varying volatility.

We explore the potential of the ESD approach in an empirical application, revisiting the construction of coincident and leading indicators for the US economy. As mentioned previously, the COVID-19 pandemic has challenged the construction of many composite indices, including those disseminated by the Philadelphia Fed. We demonstrate that the novel ESD-DFM enables us to construct the Philadelphia Fed's coincident index in the presence of the aberrant COVID-19 observations, without relying on any ad-hoc exclusions of specific time periods. We also show how the ESD-DFM methodology can be used to effectively construct the leading index, which the Philadelphia Fed has suspended in the aftermath of the pandemic. Our results show that in both cases the proposed model largely outperforms the Kalman filter-based parameter-driven DFM as well as the score-driven DFM.

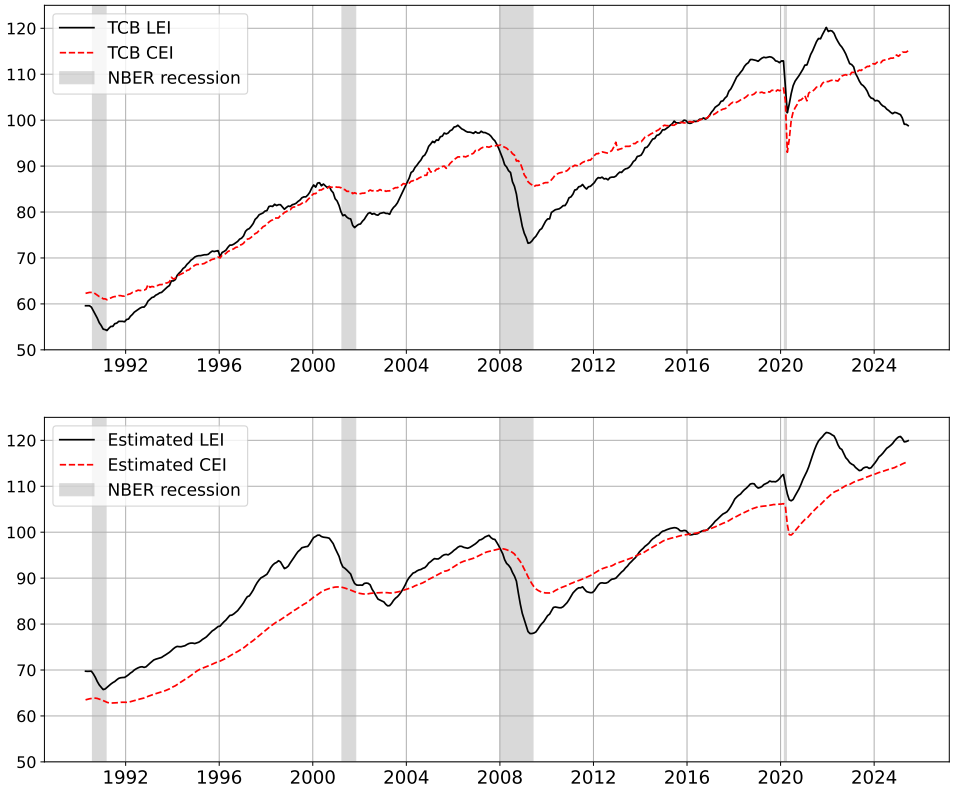


Figure 1: **TCB (top) and estimated (bottom) LEI and CEI.**

After the pandemic even non-model based composite indices, such as those published by The Conference Board (TCB), have experienced challenges. As illustrated in Figure 1,

TCB's Coincident Economic Index (TCB CEI) and Leading Economic Index (TCB LEI) have exhibited a divergent pattern over the last few years: while TCB CEI has trended upward since the pandemic, indicating healthy economic conditions, TCB LEI peaked in December 2021 and has been declining steadily since that turning point. Such divergence, persisting for more than three years at the time of writing, is unprecedented, as TCB CEI and LEI have historically moved in the same direction (obviously except during [relatively short] periods leading up to business cycle turning points). When we apply our ESD-DFM to reconstruct TCB CEI and LEI in the spirit of Stock and Watson (1989), taking into account all dynamic properties of the series, this divergence disappears. Although the estimated LEI also declines during 2022, it reaches a trough in June 2023 and exhibits an upward trend again starting from the second half of 2023 through the end of the sample in mid 2025, see the bottom panel of Figure 1.

The remainder of the paper is organized as follows. In Section 2, we illustrate a gap between the score-driven and parameter-driven DFMs using a simple example and show how our model closes this gap. In Section 3, we introduce our ESD-DFM in a general form and discusses the associated filtering and estimation procedures. In Section 4, we analyze the finite-sample performance of our estimator as well as the filtering performance in case of model misspecification. In Section 5, we apply the developed model to construct coincident and leading indicators. We conclude in Section 6.

2 Illustrative example

Before putting forward the general ESD-DFM framework, we consider a stylized example to expose the gap between parameter-driven and observation-driven DFMs and to illustrate how the ESD approach bridges this. Let $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{Nt})^\top$ denote an N-dimensional vector of observed time series, with N fixed and reasonably small. The series $\boldsymbol{y}_t$ are assumed to follow a single-factor structure:

$$\boldsymbol{y}_t = \boldsymbol{\lambda} f_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}_N, \boldsymbol{\Sigma}), \tag{1}$$

where f_t denotes the common factor and $\boldsymbol{\lambda}$ is an N-dimensional vector of loadings. The idiosyncratic components $\boldsymbol{\varepsilon}_t$ are assumed to be normally distributed with zero-mean, and are serially and cross-sectionally uncorrelated; the latter obviously implying that the covariance matrix $\boldsymbol{\Sigma}$ is diagonal. Furthermore, we follow the convention in the literature to impose the identification restriction $\frac{1}{N}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda} = 1$.

In the traditional *parameter-driven* DFM setup, the common factor f_t is assumed to follow a stationary AR(1) process, with normally distributed innovations η_t that are inde-

pendent of the innovations $\boldsymbol{\varepsilon}_t$ in the measurement equation (1):

$$f_{t+1} = \phi f_t + \eta_{t+1}, \qquad \eta_{t+1} \sim \mathcal{N}(0, \sigma_\eta^2), \tag{2}$$

with $|\phi| < 1$. Together, equations (1) and (2) form a classical Gaussian parameter-driven dynamic factor model (PD-DFM) that can be cast into a linear state-space form. The standard Kalman filter and smoother can then be applied to obtain estimates of the latent factor f_t, while the static parameters ($\boldsymbol{\lambda}$, $\boldsymbol{\Sigma}$, ϕ and σ_η^2) can be estimated using maximum likelihood (ML) via the prediction-error decomposition (Durbin and Koopman, 2012). In particular, in the steady state, the Kalman filter update for the conditional expectation of the factor, i.e., $\widehat{f}_{t+1|t} := \mathbb{E}[f_{t+1}|\mathcal{F}_t]$ with $\mathcal{F}_t$ denoting the information set at time t, is given by

$$\widehat{f}_{t+1|t} = \phi \widehat{f}_{t|t-1} + \frac{\phi p}{1 + pN} \boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}_t - \boldsymbol{\lambda} \widehat{f}_{t|t-1}), \tag{3}$$

where $p > 0$ is the steady-state predictive variance of the state. The predictive mean and variance of the data, required for the likelihood evaluation, are obtained as

$$\begin{aligned} \mathbb{E}[\boldsymbol{y}_{t+1}|\mathcal{F}_t] &= \boldsymbol{\lambda} \widehat{f}_{t+1|t}, \\ \mathbb{V}[\boldsymbol{y}_{t+1}|\mathcal{F}_t] &= p \boldsymbol{\lambda}\boldsymbol{\lambda}^\top + \boldsymbol{\Sigma}. \end{aligned}$$

Due to the presence of the innovation η_t in the factor updating equation (2), the conditional covariance matrix of the data, $\mathbb{V}[\boldsymbol{y}_{t+1}|\mathcal{F}_t]$, is non-diagonal.

By contrast, in the *score-driven* approach, the factor f_t evolves according to the updating equation

$$f_{t+1} = b f_t + a s_t,$$

with $s_t = S_t \nabla_t$, where $\nabla_t = \partial \log p_y(\boldsymbol{y}_t|f_t, \mathcal{F}_{t-1})/\partial f_t$ is the score, i.e., the gradient of the log predictive density of $\boldsymbol{y}_t$ with respect to f_t, and $S_t = I_{t|t-1}^{-1}$ is the inverse Fisher information matrix. In particular, in the Gaussian case with the same identification constraint, $\frac{1}{N}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\lambda} = 1$, the factor updating equation is given by

$$f_{t+1} = b f_t + a \frac{1}{N} \boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}_t - \boldsymbol{\lambda} f_t). \tag{4}$$

Unlike equation (2), equation (4) does not contain an exogenous innovation. In other words, the score-driven DFM (SD-DFM) given by equations (1) and (4) is *observation-*

driven, while equations (1) and (2) define a *parameter-driven* model.[3]

Notably though, equation (4) quite closely resembles the Kalman filter recursion (3) in the PD-DFM. In fact, if $b = \phi$ and $a = N\phi p/(1 + pN)$, the two filters are equivalent. However, even in that case the PD-DFM given by equations (1) and (2) and the SD-DFM given by equations (1) and (4) are not observationally equivalent: Since the factor updating equation (4) is observation-driven, the predictive variance of the factor is zero, which in turn results in the conditional covariance matrix of the data $\mathbb{V}[\boldsymbol{y}_{t+1}|\mathcal{F}_t]$ being equal to $\boldsymbol{\Sigma}$.[4] This implies that the SD-DFM, unlike the PD-DFM, misses a (quite possibly) substantial portion of the information contained in the data, as it fails to fully model the conditional second moments, i.e., the covariance terms.

This discrepancy between parameter- and observation-driven DFMs motivates us to propose a novel class of score-driven models, for which the observational equivalence with PD-DFMs can be established. We formally introduce this class of *extended* score-driven DFMs (ESD-DFMs) in the next section. The label *extended* score-driven is based on the fact that the model extends the SD-DFM by augmenting the factor updating equation with a contemporaneous score. Hence, in the current illustrative example of a single-factor model, the ESD updating equation for the factor reads as

$$f_{t+1} = bf_t + as_t + cs_{t+1}.$$

In the Gaussian case, the factor updating equation thus becomes

$$f_{t+1} = bf_t + a\frac{1}{N}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{y}_t - \boldsymbol{\lambda} f_t\right) + c\frac{1}{N}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{y}_{t+1} - \boldsymbol{\lambda} f_{t+1}\right). \tag{5}$$

An important difference between the ESD-DFM updating equation in (5) and the conventional SD-DFM update in equation (4) is that the last term on the right-hand side of (5) is not measurable with respect to the filtration $\mathcal{F}_t$. Therefore, the predictive factor is no longer exact and the factor recursion mechanism naturally splits into two parts as in the Kalman filter: updating and prediction steps. The prediction step, conditional on the information at time t, is given by

$$f_{t+1|t} := \mathbb{E}[f_{t+1}|\mathcal{F}_t] = bf_t + a\frac{1}{N}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{y}_t - \boldsymbol{\lambda} f_t\right).$$

[3] Another prominent example of an observation-driven model is the generalised autoregressive conditional heteroskedasticity (GARCH) model, which is widely used to model time-varying volatility in financial time series. GARCH models are often contrasted with parameter-driven stochastic volatility (SV) models, in which volatility dynamics are driven by exogenous innovations.

[4] This is in contrast to the univariate Gaussian SD and PD location models, which are both equivalent to a (restricted) ARMA(1,1) model (Harvey, 1990; Artemova et al., 2022).

For the updating step, i.e., conditioning on $\mathcal{F}_{t+1}$, by solving equation (5) for f_{t+1} we obtain

$$\begin{aligned}\mathbb{E}[f_{t+1}|\mathcal{F}_{t+1}] = f_{t+1} &= \frac{1}{1+c} f_{t+1|t} + \frac{c}{1+c}\frac{1}{N}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{y}_{t+1} \\ &= f_{t+1|t} + \frac{c}{1+c}\frac{1}{N}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{y}_{t+1} - \boldsymbol{\lambda} f_{t+1|t}\right),\end{aligned} \tag{6}$$

where the first equality holds since the factor f_{t+1} is $\mathcal{F}_{t+1}$-measurable. The parameter c determines the weight the updating step puts on the prediction $f_{t+1|t}$ relative to the information from the newly incoming observation $\boldsymbol{y}_{t+1}$. Specifically, smaller values of c assign more weight to the prediction, while larger values of c increase the influence of the new observation in updating the factor estimate.

It is now straightforward to show that in the ESD-DFM context, the predictive mean and variance of the data are given by

$$\begin{aligned}\mathbb{E}[\boldsymbol{y}_{t+1}|\mathcal{F}_t] &= \boldsymbol{\lambda} f_{t+1|t}, \\ \mathbb{V}[\boldsymbol{y}_{t+1}|\mathcal{F}_t] &= \frac{c^2+2c}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top + \boldsymbol{\Sigma}.\end{aligned}$$

Thus, unlike the SD-DFM, the predictive conditional variance of the data is non-diagonal when $c \neq 0$. Furthermore, as we show in Appendix A.1, the ESD-DFM with Gaussian density given by equations (1) and (5), with $c = \frac{a}{b-a}$, is observationally equivalent to the PD-DFM defined by equations (1) and (2). We emphasize again that this is in contrast to the conventional SD-DFM, for which this observational equivalence does not hold. Put differently, the ESD-DFM bridges the PD- and SD-DFMs: $c = 0$ yields the SD-DFM setup, while the Gaussian model together with $c = \frac{a}{b-a}$ is observationally equivalent to the Gaussian PD-DFM.

What is more, the ESD-DFM can be more flexible than both the PD- and SD-DFMs. As we show in Appendix A.2, unlike the traditional PD-DFM, the ESD-DFM can generate non-diagonal autocovariance in the vector moving average (MA) component of the data. In fact, combining equations (1) and (5) we can express the observation $\boldsymbol{y}_{t+1}$ generated by the ESD-DFM as

$$\begin{aligned}\boldsymbol{y}_{t+1} &= b\boldsymbol{y}_t + a\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_t + c\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_{t+1} + \boldsymbol{\varepsilon}_{t+1} - b\boldsymbol{\varepsilon}_t \\ &= b\boldsymbol{y}_t + \left(\boldsymbol{I} + c\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{\varepsilon}_{t+1} - b\left(\boldsymbol{I} - \frac{a}{b}\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{\varepsilon}_t.\end{aligned}$$

Given this representation, we can show (see Appendix A.2 for details) that

$$\mathbb{C}ov(\boldsymbol{y}_{t+1} - b\boldsymbol{y}_t, \boldsymbol{y}_t - b\boldsymbol{y}_{t-1}) = -b\boldsymbol{\Sigma} + (a - bc + ac)\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top. \tag{7}$$

For the PD-DFM with $c = \frac{a}{b-a}$, the second term in equation (7) is equal to zero, resulting in a diagonal autocovariance matrix, with the sign of the diagonal elements being opposite from the persistence parameter $b = \phi$ in the VAR part. The latter is analogous to the univariate AR(1)-plus-noise model being a restricted version of an ARMA(1,1) process. Whenever $c \neq \frac{a}{b-a}$, the ESD-DFM is able to generate a non-diagonal autocovariance term of the vector MA term with unrestricted signs on its diagonal elements, resulting in richer dynamics.

	PD-$\mathcal{N}$	SD-$\mathcal{N}$	ESD-$\mathcal{N}$
Loglik	-3672.75	-3724.64	**-3658.07**
c	0.37	–	0.86

Table 1: **Model fit comparison.** We consider the four PHI CEI constituents with data before 2020. We report the log-likelihood value (Loglik) for PD-, SD-, and ESD-DFM with Gaussian distributions as well as the corresponding c parameters.

We conclude this section with a preview of one of the empirical applications, designed to illustrate the practical implications of the modeling choices discussed above. Specifically, we use four monthly macroeconomic series that enter the construction of the Philadelphia Fed Coincident Economic Index (PHI CEI); see Section 5 for further details. Using data up to 2019, we estimate the three DFMs considered in this section. The log-likelihood values in Table 1 show that the ESD-DFM achieves the best fit, followed by the PD-DFM and then the SD-DFM. This ranking indicates that the ESD-DFM can capture more information from the data than the SD model, while extending the PD-DFM through the additional flexibility introduced by the c parameter. The estimated parameter c in the ESD-DFM is found to be nonzero and exceeds the value implied by the restriction in the parameter-driven model, underscoring the generality of the ESD-DFM framework. Furthermore, the ESD-DFM model can easily be made even more general by allowing for non-Gaussian innovations and nonlinear updates with minimal computational cost, unlike PD-DFMs, where such extensions are typically computationally demanding. The next sections introduce these generalizations and further demonstrate the flexibility and empirical relevance of the framework.

3 ESD framework: model, filter, and estimation

3.1 Model specification

In this section, we develop the general framework of the extended score-driven dynamic factor model (ESD-DFM), which accommodates multiple factors, non-linear factor dynamics, serially correlated idiosyncratic components, non-Gaussian error distributions, and conditional heteroskedasticity.

Let $\boldsymbol{y}_t = (y_{1,t}, \dots y_{N,t})^\top$ be an N-dimensional vector of time series variables indexed by time t. We consider a dynamic factor model for $\boldsymbol{y}_t$, with serially correlated and conditionally heteroskedastic idiosyncratic components $\boldsymbol{\varepsilon}_t$, i.e.,

$$\boldsymbol{y}_t = \boldsymbol{\Lambda}_0 \boldsymbol{f}_t + \boldsymbol{\Lambda}_1 \boldsymbol{f}_{t-1} + \ldots + \boldsymbol{\Lambda}_m \boldsymbol{f}_{t-m} + \boldsymbol{\varepsilon}_t, \tag{8}$$

$$\boldsymbol{\varepsilon}_t = \boldsymbol{P}_1 \boldsymbol{\varepsilon}_{t-1} + \ldots + \boldsymbol{P}_p \boldsymbol{\varepsilon}_{t-p} + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \sim p_u(\boldsymbol{u}_t, \boldsymbol{\Sigma}_t; \boldsymbol{\nu}), \tag{9}$$

where $\boldsymbol{f}_t$ is an $r \times 1$ vector of common factors, each $\boldsymbol{\Lambda}_j$ is an $N \times r$ matrix of loadings on $\boldsymbol{f}_{t-j}$ for $j = 0, \dots, m$, and each $\boldsymbol{P}_j$ is an $N \times N$ diagonal autoregressive matrix for $j = 1, \dots, p$. The innovations $\boldsymbol{u}_t$ are assumed to be independently distributed with density $p_u(\boldsymbol{u}_t, \boldsymbol{\Sigma}_t; \boldsymbol{\nu})$, with mean zero, time-varying diagonal conditional scale matrix $\boldsymbol{\Sigma}_t$, and additional static shape parameters $\boldsymbol{\nu}$. Equations (8) and (9) can also be expressed as follows:

$$\boldsymbol{y}_t = \boldsymbol{\Lambda}(L)\boldsymbol{f}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{P}(L)\boldsymbol{\varepsilon}_t = \boldsymbol{u}_t,$$

where $\boldsymbol{\Lambda}(L) = \boldsymbol{\Lambda}_0 + \boldsymbol{\Lambda}_1 L + \ldots + \boldsymbol{\Lambda}_m L^m$ and $\boldsymbol{P}(L) = \boldsymbol{I} - \boldsymbol{P}_1 L - \ldots - \boldsymbol{P}_p L^p$ with the lag operator L defined as $L^k z_t = z_{t-k}$ for any integer k. The lag structure of the measurement equation as well as allowing for serially correlated idiosyncratic components are common in the DFM literature (see, e.g., Stock and Watson, 1989). However, we further allow for more general innovation distributions and for conditional heteroskedasticity.

We model the factor dynamics using an observation-driven framework. In particular, the factors are assumed to follow an extended score-driven mechanism:

$$\boldsymbol{f}_{t+1} = \boldsymbol{B}(L)\boldsymbol{f}_t + \boldsymbol{A}\boldsymbol{s}_t^{\boldsymbol{f}} + \boldsymbol{C}\tilde{\boldsymbol{s}}_{t+1}^{\boldsymbol{f}}, \tag{10}$$

where $\boldsymbol{B}(L) = \boldsymbol{B}_0 + \boldsymbol{B}_1 L + \ldots + \boldsymbol{B}_q L^q$; and $\boldsymbol{A}$, $\boldsymbol{C}$, and $\boldsymbol{B}_i$ for $i = 0, \dots, q$ are $r \times r$ matrices. Following the literature on score-driven models, we set the score $\boldsymbol{s}_t^{\boldsymbol{f}}$ equal to the score of the conditional likelihood of the data $\boldsymbol{y}_t$, as implied by a prespecified distribution of the error term $\boldsymbol{u}_t$. This paper considers the Gaussian and Student's t distributions for the error term $\boldsymbol{u}_t$, with the corresponding score expressions discussed below (see also Section 2 in

Artemova, 2025). The final term $\tilde{\boldsymbol{s}}^{\boldsymbol{f}}_{t+1}$ is the (quasi) score evaluated at time $t+1$ which we discuss in detail below.

The main novelty of our ESD-DFM lies in the inclusion of the scores at both times t and $t+1$ in the updating equation (10). It is exactly this feature that builds the bridge between the conventional observation-driven and parameter-driven DFMs. Specifically, when $\boldsymbol{C} = \boldsymbol{0}$ our model collapses to the SD-DFM as considered in, e.g., Artemova (2025). In turn, when $\boldsymbol{C} \neq \boldsymbol{0}$ the factors $\boldsymbol{f}_{t+1}$ are no longer predetermined at time t, but instead become stochastic as typically the case in parameter-driven DFMs. The idea of incorporating the time $t+1$ innovations is conceptually similar to the real-time GARCH model of Smetanina (2017), which bridges the gap between GARCH and stochastic volatility (SV) models. Similarly, this paper aims to close the gap between Gaussian score-driven and parameter-driven DFMs, while extending the framework to accommodate deviations from Gaussianity and the presence of heteroskedasticity.

For the analytic tractability of the likelihood function, we adopt the Gaussian score for $\tilde{\boldsymbol{s}}^{\boldsymbol{f}}_{t+1}$, such that

$$\tilde{\boldsymbol{s}}^{\boldsymbol{f}}_{t+1} := \boldsymbol{S}_{t+1}\boldsymbol{\Lambda}_0^\top \boldsymbol{\Sigma}_{t+1}^{-1}\boldsymbol{P}(L)[\boldsymbol{y}_{t+1} - \boldsymbol{\Lambda}(L)\boldsymbol{f}_{t+1}] = \boldsymbol{S}_{t+1}\boldsymbol{\Lambda}_0^\top \boldsymbol{\Sigma}_{t+1}^{-1}\boldsymbol{u}_{t+1},$$

where $\boldsymbol{S}_t := (\boldsymbol{\Lambda}_0^\top \boldsymbol{\Sigma}_t^{-1}\boldsymbol{\Lambda}_0)^{-1}$ is the scaling matrix. Thus, effectively we include the time $t+1$ observation innovations in the factor updating mechanism (10). This is in contrast to the PD-DFM, where the innovations in the factor updating and measurement equations are independent and represent different sources of uncertainty. By utilizing the innovation terms from the measurement equation in the factor dynamics, the ESD-DFM remains within the observation-driven framework.

Moreover, the ESD-DFM implies a non-diagonal conditional covariance matrix $\mathbb{V}ar[\boldsymbol{y}_t|\mathcal{F}_{t-1}]$, similar to the PD-DFM and in contrast to the SD-DFM, where this matrix is diagonal by construction. In Appendix A.1, we show that in the special case of Gaussian innovations, the ESD-DFM nests the PD-DFM. Therefore, the ESD-DFM preserves the flexibility and computational advantages of the score-driven framework while capturing all dynamic features of the data.

We also allow the conditional scale matrix $\boldsymbol{\Sigma}_t$ to be time-varying. For parsimony, we impose a factor structure of the form $\boldsymbol{\Sigma}_t = h_t^2\boldsymbol{\Sigma}$ with a diagonal matrix $\boldsymbol{\Sigma}$ and the volatility factor h_t evolving according to a GARCH-type specification[5]:

$$h_{t+1}^2 = w + \alpha s_t^h + \gamma h_t^2, \tag{11}$$

[5]The model can potentially be extended to include a multivariate volatility factor at the cost of increased notational complexity and larger number of parameters.

where s_t^h is the scaled score corresponding to the volatility process. With the Gaussian score, equation (11) becomes the GARCH(1,1)-type specification for the common scale factor, while the Student's t score produces a more robust update for the scale parameter (see Blasques et al., 2022). Naturally, with $w = 1$ and $\alpha = \gamma = 0$, we retrieve the exact dynamic factor model without conditional heteroskedasticity.

We note that the three scores $\boldsymbol{s}_t^{\boldsymbol{f}}$, $\tilde{\boldsymbol{s}}_{t+1}^{\boldsymbol{f}}$ and s_t^h do not all have to correspond to the same density function $p_u(\boldsymbol{u}_t, \boldsymbol{\Sigma}_t; \boldsymbol{\nu})$. In such cases, our model falls in the (extended) quasi score-driven framework, as introduced in Blasques et al. (2023). In a single-factor model without autocorrelation and under Gaussian scores, the updating equation is given by equation (5) presented in Section 2. Below we present an example of the model based on the Student's t distribution with the GARCH(1,1)-type update for the common scale factor, which we refer to as the ESD-t DFM.

EXAMPLE 1 (The Student's t ESD-DFM)**.** *For the Student's t model with $m = q = p = 0$, the scaled (quasi) scores are of the form $\boldsymbol{s}_t^{\boldsymbol{f}} = \frac{1}{W_t}\boldsymbol{S}\boldsymbol{\Lambda}_0^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{u}_t$, $\tilde{\boldsymbol{s}}_{t+1}^{\boldsymbol{f}} = \boldsymbol{S}\boldsymbol{\Lambda}_0^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{u}_{t+1}$, and $s_t^h = \frac{1}{N}\boldsymbol{u}_t^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{u}_t - h_t^2$; see derivations in Koopman et al. (2017); Artemova (2025). Then, the updating equations for the factors and scale parameters are*

$$\begin{aligned}
\boldsymbol{f}_{t+1} &= \boldsymbol{B}\boldsymbol{f}_t + \boldsymbol{A}\frac{1}{W_t}\boldsymbol{S}\boldsymbol{\Lambda}_0^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{u}_t + \boldsymbol{C}\boldsymbol{S}\boldsymbol{\Lambda}_0^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{u}_{t+1}, \\
h_{t+1}^2 &= w + \alpha\frac{1}{N}\boldsymbol{u}_t^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{u}_t + (\gamma - \alpha)h_t^2,
\end{aligned}$$

where $\boldsymbol{S} := (\boldsymbol{\Lambda}_0^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}_0)^{-1}$ and $W_t := \frac{\nu}{N+\nu+2}(1 + \boldsymbol{u}_t^\top \boldsymbol{\Sigma}_t^{-1}\boldsymbol{u}_t/\nu)$.

The weight W_t in Example 1 acts as a robustness mechanism in the factor update, dampening the impact of extreme observations. As the degrees of freedom parameter $\nu \to \infty$, $W_t \to 1$ and we recover the Gaussian case. In the empirical analysis, we demonstrate that incorporating the robustness feature is particularly important in macroeconomic DFM applications, especially when the sample period includes the COVID-19 pandemic.

3.2 Filter

As mentioned before, when $\boldsymbol{C} \neq \boldsymbol{0}$ in equation (10), the factors $\boldsymbol{f}_{t+1}$ are not predetermined at time t. However, once information at time $t + 1$ (in particular the observation $\boldsymbol{y}_{t+1}$) becomes available, $\boldsymbol{f}_{t+1}$ can be fully recovered. In other words, the factors $\boldsymbol{f}_{t+1}$ are $\mathcal{F}_{t+1}$-measurable but not $\mathcal{F}_t$-measurable. Therefore, it is convenient and helpful to split the filtering recursion given by equation (10) into two steps: prediction and updating steps – analogously to the Kalman filter.

In the ESD-DFM with a single volatility factor and with the Gaussian score at time $t+1$, i.e., $\tilde{\boldsymbol{s}}^{\boldsymbol{f}}_{t+1} = \boldsymbol{S}\boldsymbol{\Lambda}_0^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{P}(L)\left[\boldsymbol{y}_{t+1} - \boldsymbol{\Lambda}(L)\boldsymbol{f}_{t+1}\right]$, the prediction step reads

$$\boldsymbol{f}_{t+1|t} := \mathbb{E}[\boldsymbol{f}_{t+1}|\mathcal{F}_t] = \boldsymbol{B}(L)\boldsymbol{f}_t + \boldsymbol{A}\boldsymbol{s}^{\boldsymbol{f}}_t, \tag{12}$$

since $\mathbb{E}\left[\tilde{\boldsymbol{s}}^{\boldsymbol{f}}_{t+1}|\mathcal{F}_t\right] = 0$. Note that the right-hand side of equation (12) is given in terms of $\boldsymbol{f}_t$ since $\mathbb{E}[\boldsymbol{f}_t|\mathcal{F}_t] = \boldsymbol{f}_t$. Not surprisingly, the prediction step coincides with the SD-DFM update of the factor. However, in contrast to the conventional score-driven models, in the ESD setting the conditional variance of the predicted factor is not zero and is given by

$$\begin{aligned}\mathbb{V}[\boldsymbol{f}_{t+1}|\mathcal{F}_t] = \boldsymbol{C}\mathbb{E}\left[\tilde{\boldsymbol{s}}^{\boldsymbol{f}}_{t+1}\tilde{\boldsymbol{s}}^{\boldsymbol{f}\top}_{t+1}|\mathcal{F}_t\right]\boldsymbol{C}^\top &= \boldsymbol{C}\boldsymbol{S}\boldsymbol{\Lambda}_0^\top\boldsymbol{\Sigma}_{t+1}^{-1}\boldsymbol{\Lambda}_0\boldsymbol{S}\boldsymbol{C}^\top \\ &= h^2_{t+1}\boldsymbol{C}\boldsymbol{S}\boldsymbol{C}^\top.\end{aligned}$$

As we have discussed in the previous subsection, the non-zero variance matrix of the predicted factors bridges the gap between the SD-DFM and PD-DFM. This also allows us to construct confidence intervals around the predicted factors, analogously to the state-space framework.

For the updating step, utilizing the fact that $\boldsymbol{f}_{t+1} \equiv \mathbb{E}[\boldsymbol{f}_{t+1}|\mathcal{F}_{t+1}]$, it follows that

$$\begin{aligned}\boldsymbol{f}_{t+1} &= \boldsymbol{f}_{t+1|t} + \boldsymbol{C}\boldsymbol{S}\boldsymbol{\Lambda}_0^\top\boldsymbol{\Sigma}^{-1}\mathbb{E}\left[\boldsymbol{P}(L)(\boldsymbol{y}_{t+1} - \boldsymbol{\Lambda}(L)\boldsymbol{f}_{t+1})|\mathcal{F}_{t+1}\right], \\ &= \boldsymbol{f}_{t+1|t} + \boldsymbol{C}\boldsymbol{S}\boldsymbol{\Lambda}_0^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_{t+1} - \boldsymbol{\Lambda}(L)\boldsymbol{f}_{t+1}) - \boldsymbol{C}\boldsymbol{\Lambda}_0^\top\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{P}}(L)(\boldsymbol{y}_t - \boldsymbol{\Lambda}(L)\boldsymbol{f}_t), \\ &= \boldsymbol{f}_{t+1|t} - \boldsymbol{C}\boldsymbol{f}_{t+1} + \boldsymbol{C}\boldsymbol{f}_{t+1|t} \\ &\quad + \boldsymbol{C}\boldsymbol{S}\boldsymbol{\Lambda}_0^\top\boldsymbol{\Sigma}^{-1}\underbrace{\left(\boldsymbol{y}_{t+1} - \boldsymbol{\Lambda}_0\boldsymbol{f}_{t+1|t} - \tilde{\boldsymbol{\Lambda}}(L)\boldsymbol{f}_t - \tilde{\boldsymbol{P}}(L)\left(\boldsymbol{y}_t - \boldsymbol{\Lambda}(L)\boldsymbol{f}_t\right)\right)}_{=:\boldsymbol{e}_{t+1}},\end{aligned}$$

where $\tilde{\boldsymbol{P}}(L) = \boldsymbol{P}_1 + \ldots + \boldsymbol{P}_p L^{p-1}$, $\tilde{\boldsymbol{\Lambda}}(L) = \boldsymbol{\Lambda}_1 + \ldots + \boldsymbol{\Lambda}_m L^{m-1}$, and $\boldsymbol{e}_{t+1}$ is the one-step-ahead prediction error of the data. Therefore, the updating step can be written as

$$\boldsymbol{f}_{t+1} = \boldsymbol{f}_{t+1|t} + (\boldsymbol{I} + \boldsymbol{C})^{-1}\boldsymbol{C}\boldsymbol{S}\boldsymbol{\Lambda}_0^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{e}_{t+1}. \tag{13}$$

The ESD filter alternates between the prediction and updating steps given by equations (12) and (13), respectively. The intuition is analogous to the Kalman filter: once new information at time $t+1$ becomes available, the prediction is updated based upon the realized prediction error. Typically, in the score-driven literature the distinction between prediction and updating steps is not made as these steps coincide. However, in our ESD set-up, this is not the case since the factor $\boldsymbol{f}_{t+1}$ is not $\mathcal{F}_t$-measurable. The importance of this split for observation-driven models has also been recently highlighted in (Lange et al., 2024) for their implicit score-driven filter.

3.3 Estimation

We can estimate the ESD-DFM parameters by (quasi) maximum likelihood (ML) using the prediction error decomposition, as is standard for both state-space and observation-driven models. Specifically, the ML estimator is

$$\hat{\boldsymbol{\theta}}_T = \arg\max_{\boldsymbol{\theta}\in\Theta} L_T(\boldsymbol{\theta}),$$

where $\boldsymbol{\theta}$ is a vector collecting all model parameters, Θ indicates the parameter space, and $L_T(\boldsymbol{\theta})$ denotes the log-likelihood function given by

$$L_T(\boldsymbol{\theta}) = \sum_{t=2}^{T} \left(-\frac{1}{2}\log|\boldsymbol{\Omega}_t| + \log g\left(\boldsymbol{e}_t^\top \boldsymbol{\Omega}_t^{-1}\boldsymbol{e}_t\right)\right),$$

where $\boldsymbol{e}_t = \boldsymbol{y}_t - \mathbb{E}[\boldsymbol{y}_t|\mathcal{F}_{t-1}]$ denotes the one-step-ahead prediction error, $\boldsymbol{\Omega}_t = \mathbb{V}(\boldsymbol{e}_t|\mathcal{F}_{t-1}) = \mathbb{V}(\boldsymbol{y}_t|\mathcal{F}_{t-1})$ is the conditional variance matrix, and $g(\cdot)$ is the density generator of $p_u(\boldsymbol{u}_t, \boldsymbol{\Sigma}_t; \boldsymbol{\nu})$. The asymptotic properties of the estimator largely follow from the theory developed for score-driven factor models (see Artemova, 2025).

For both the asymptotic analysis and numerical optimization, parameter identification must be ensured. Identification issues in dynamic factor models have been widely studied. In particular, Bai and Wang (2015) address identification in parameter-driven specifications, whereas Artemova (2025) investigates the corresponding issues in score-driven models. Identification of the common component $\boldsymbol{\Lambda}(L)\boldsymbol{f}_t$ and the covariance matrix can be achieved by assuming that $\boldsymbol{\Sigma}$ is diagonal. Under this assumption, all information in the off-diagonal elements of the unconditional covariance matrix of the data is attributed to the factors (Anderson et al., 1956).

Assumption 1. *The scale matrix $\boldsymbol{\Sigma}$ is diagonal with elements $0 < \sigma_i^2 < \infty$ for all $i = 1, \ldots, N$.*

A further challenge in factor models is the problem of rotational indeterminacy: the factors and loadings can be rotated without changing the likelihood, which complicates identification. Artemova (2025) studies the identification problem in models without lags of the factors in the observation equation. Therefore, we first establish the conditions required for identification in the extended score-driven dynamic factor model.

Assumption 2. $\boldsymbol{S}_0 = \frac{1}{N}\boldsymbol{\Lambda}_0^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}_0 \succ 0$.

Assumption 3. *Matrix $\boldsymbol{\Lambda}_0$ is of rank r.*

Lemma 1. *Let assumptions 1–3 hold, then factors and loadings in the (extended) score-driven dynamic factor model* (8)–(10) *are identified up to a $q \times q$ rotation.*

To resolve the rotational indeterminacy, further restrictions on the loadings matrix are typically imposed – such as those proposed by Bai and Li (2012) and adopted in Artemova (2025) for the case $m = 0$, and by Bai and Wang (2015) for $m \geq 1$. In the context of score-driven models, an alternative strategy involves using a power of the inverse Fisher information matrix different from one, as suggested by Buccheri et al. (2024), which also aids in achieving identification. For example, we can impose the following restrictions

1. $\frac{1}{N}\boldsymbol{\Lambda}_0^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}_0 = \boldsymbol{I}_r$ and $\mathbb{C}ov(\boldsymbol{f}_t)$ is diagonal with distinct elements on the diagonal;
2. $\boldsymbol{\Lambda}_0 = [\boldsymbol{\Lambda}_{0,1}, \boldsymbol{\Lambda}_{0,2}]^\top$ with $\boldsymbol{\Lambda}_{0,1} = \boldsymbol{I}_r$.

Both restrictions are equivalent in the case of a single factor ($r = 1$). For $r \geq 2$, the first condition is invariant to the ordering of the series; however, it requires the factors to be uncorrelated, implying diagonal restrictions on $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ in equation (10) and thereby excluding spillover effects. In contrast, the second restriction allows for spillover effects but requires a specific ordering of the series. In the remainder of the paper, we assume that one of the identification conditions is imposed. In the empirical application, we adopt the first condition, which ensures that $\boldsymbol{\Lambda}$ and $\mathbb{C}ov(\boldsymbol{f}_t)$ are identifiable.

In terms of parameters, the key difference between the SD and ESD models is $\boldsymbol{C}$, the parameter for the contemporaneous score in the factor update equation (10). In what follows, we establish identification of this parameter conditional on the remaining parameters, and we prove joint identification for Example 1. The proofs for other specifications are analogous but involve cumbersome expressions, and are omitted for brevity.

Assumption 4. *$\boldsymbol{C}$ is diagonal with positive elements, i.e., $\boldsymbol{C} = \mathrm{diag}(c_1, \ldots, c_r)$ and $c_s > 0$ for all $s = 1, \ldots, r$.*

Assumption 5. $\mathbb{E}\|\boldsymbol{y}_t\|^2 < \infty$ *and* $\mathbb{E}\|\boldsymbol{f}_t(\boldsymbol{\theta}_0)\|^2 < \infty$.

Lemma 2. *Under Assumptions 4 and 5 and for fixed values of all other parameters $\boldsymbol{\theta}_{-\boldsymbol{C}}$, the parameter $\boldsymbol{C}$ is identified.*

4 Monte Carlo simulations

In this section, we analyze the finite-sample properties of the ML estimator under correct specification, as well as the ESD filtering performance in the presence of model misspecification. For this, we consider two simulation setups: one with the ESD-DFM as the data generating process (DGP), and another with the parameter-driven model as the DGP.

4.1 Finite sample properties ML estimator

In this simulation, the DGP is a Gaussian ESD-DFM. Specifically, we consider a single factor specification, with one lag of the factor in the observation equation, and first-order dynamics for both the common factor as well as the idiosyncratic components:

$$\begin{aligned}
\boldsymbol{y}_t &= \boldsymbol{\lambda}_0 f_t + \boldsymbol{\lambda}_1 f_{t-1} + \boldsymbol{\varepsilon}_t, \\
\boldsymbol{\varepsilon}_t &= \boldsymbol{P}\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{u}_t, \\
f_{t+1} &= a s_t + c s_{t+1} + b f_t,
\end{aligned}$$

where $s_t = (\boldsymbol{\lambda}_0^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda}_0)^{-1}\boldsymbol{\lambda}_0^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{u}_t$ and $\boldsymbol{u}_t = (\boldsymbol{y}_t - \boldsymbol{\lambda}_0 f_t - \boldsymbol{\lambda}_1 f_{t-1}) - \boldsymbol{P}(\boldsymbol{y}_{t-1} - \boldsymbol{\lambda}_0 f_{t-1} - \boldsymbol{\lambda}_1 f_{t-2})$. The values of the parameters were chosen such that they align with the empirical application. Specifically, we consider $N = 4$ and set $a = 0.5$, $c = 1$, $b = 0.9$, $\boldsymbol{\lambda}_1 = (0, -0.1, -0.3, -0.4)^\top$, $\boldsymbol{P} = \text{diag}(-0.1, 0.1, 0.05, -0.25)$, and $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2)$ with $\sigma_i^2 = 0.05 + 0.15 \times (i-1)$ for $i = 1, \ldots, 4$. The vector $\boldsymbol{\lambda}_0$ is initially set to $(0.2, 0.2, 0.4, 0.5)^\top$ and is then rotated to satisfy the normalization $\frac{1}{N}\sum_{i=1}^N \lambda_{0,i}^2 = 1$.

We consider two sample sizes, $T = 500$ and $T = 1000$. In all simulations, the first 200 observations are discarded as burn-in to mitigate initialization effects. The results are based on 1000 Monte Carlo replications.

The bias estimates and corresponding standard errors are reported in Table 2. The results indicate that the ML estimator performs well in finite samples. For all parameters the biases are small and decrease (both in terms of magnitude as well as in variation) when the sample size T becomes larger.

4.2 Model fit and filtering

In this simulation, we use several PD-DFMs as data generating processes (DGPs) and compare the in- and out-of-sample performance of the PD and (E)SD DFMs and the associated filters. Specifically, we simulate data according to

$$\begin{aligned}
\boldsymbol{y}_t &= \boldsymbol{\lambda}_0 f_t + \boldsymbol{\varepsilon}_t, \\
\boldsymbol{\varepsilon}_t &= \boldsymbol{P}\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{u}_t, \\
f_t &= \phi f_{t-1} + \kappa \eta_t,
\end{aligned}$$

where the innovations $\boldsymbol{u}_t$ and η_t are independent and follow either Gaussian or Student's t distributions: $\boldsymbol{u}_t \sim \mathcal{N}(\boldsymbol{0}_N, \boldsymbol{\Sigma})$ and $\eta_t \sim \mathcal{N}(0,1)$, or $\boldsymbol{u}_t \sim t_3(\boldsymbol{0}_N, \boldsymbol{\Sigma})$ and $\eta_t \sim t_6(0,1)$, respectively.

Throughout the simulations, we set $N = 4$ and $T = 2 \times 500$, with a burn-in period of 200

	a	b	c	$\lambda_{0,i}$	$\lambda_{1,i}$	σ_i^2	p_i
$T = 500$							
	0.002	-0.005	0.010				
	(0.105)	(0.025)	(0.079)				
$i = 1$				0.000	–	-0.001	-0.002
				(0.035)	–	(0.005)	(0.058)
$i = 2$				0.003	-0.004	-0.002	-0.003
				(0.053)	(0.061)	(0.014)	(0.047)
$i = 3$				-0.000	0.002	-0.001	-0.006
				(0.059)	(0.085)	(0.028)	(0.049)
$i = 4$				-0.005	0.006	-0.004	0.002
				(0.052)	(0.094)	(0.041)	(0.050)
$T = 1000$							
	0.000	-0.002	0.007				
	(0.073)	(0.016)	(0.056)				
$i = 1$				0.001	–	-0.000	-0.002
				(0.024)	–	(0.004)	(0.040)
$i = 2$				-0.001	0.001	-0.000	-0.001
				(0.037)	(0.043)	(0.010)	(0.034)
$i = 3$				-0.000	0.002	-0.000	-0.002
				(0.042)	(0.060)	(0.019)	(0.037)
$i = 4$				-0.002	0.002	-0.001	-0.001
				(0.037)	(0.066)	(0.029)	(0.035)

Table 2: **Monte Carlo simulation results.** The table reports the biases and standard errors for the static parameters with $i = 1, \ldots, 4$ and two different sample sizes $T = \{500, 1000\}$.

observations to reduce the influence of a starting point. The first 500 observations are used for parameter estimation, and the second 500 for out-of-sample evaluation. The remaining parameters again are chosen to align with the empirical application. Specifically, we set $\kappa = 0.5$, $\phi = 0.9$, and $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2)$ with $\sigma_i^2 = 0.2 + 0.2 \times (i-1)$ for $i = 1, \ldots, 4$. The factor loadings are given by $\boldsymbol{\lambda}_0 = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)^\top$, where $\lambda_i = 1 - 0.25 \times (i-1)$ for $i = 1, \ldots, 4$. For each DGP, we consider two cases: one without autocorrelation in the idiosyncratic components ($\boldsymbol{P} = \boldsymbol{0}$) and one with moderate autocorrelation, specified as $\boldsymbol{P} = \text{diag}(p_1, p_2, p_3, p_4)$ with $p_i = -0.1 \times i$ for $i = 1, \ldots, 4$. Notably, except for the Gaussian DGP, both the PD and ESD models are misspecified, while the SD model is always misspecified.

Table 3 reports results for the in-sample fit in terms of log-likelihood and consistent Akaike information criterion (cAIC), both evaluated at the estimated parameters. The corresponding out-of-sample results, based on the logarithmic scoring rule (LSR), mean squared error (MSE) of the factor and MSE of the first observed series, are presented in Table 4. These results are based on 300 Monte Carlo replications.

Concerning the in-sample performance, we find that the ESD models always achieve higher log-likelihoods than the PD-$\mathcal{N}$ model. As expected under the Gaussian DGP 1, this

	PD-$\mathcal{N}$	SD-$\mathcal{N}$	ESD-$\mathcal{N}$	SD-t	ESD-t
DGP 1: $\mathcal{N}$, $\boldsymbol{P} = \mathbf{0}$					
Loglike	-3186.44	-3305.21	-3185.97	-3302.90	**-3185.76**
cAIC	**6440.69**	6678.23	6447.29	6681.14	6454.38
DGP 2: t, $\boldsymbol{P} = \mathbf{0}$					
Loglike	-4537.09	-4604.90	-4534.68	-4176.20	**-4055.71**
cAIC	9141.99	9277.60	9144.71	8427.74	**8194.30**
DGP 3: $\mathcal{N}$, $\boldsymbol{P} \neq \mathbf{0}$					
Loglike	-3205.15	-3341.20	-3205.15	-3338.55	**-3204.96**
cAIC	**6508.25**	6780.35	6515.78	6782.58	6522.93
DGP 4: t, $\boldsymbol{P} \neq \mathbf{0}$					
Loglike	-4538.52	-4618.84	-4536.51	-4218.76	**-4076.02**
cAIC	9174.98	9335.62	9178.49	8542.99	**8265.05**

Table 3: **In-sample Monte Carlo results.**

difference is marginal, since the PD-$\mathcal{N}$ is correctly specified, and the ESD-$\mathcal{N}$ nests PD-$\mathcal{N}$. Because the PD-$\mathcal{N}$ model is correctly specified and more parsimonious, the information criterion naturally favors it. This pattern is also reflected in the out-of-sample results, where the PD-$\mathcal{N}$ is always performing best in the Gaussian settings. Nevertheless, the ESD-$\mathcal{N}$ and ESD-t models are very competitive to the PD-$\mathcal{N}$ model when the DGP is Gaussian, both in- and out-of-sample, demonstrating the ability of the ESD modeling approach to effectively approximate parameter-driven DGPs.

When the DGP features Student's t innovations (DGP 2 and DGP 4), the results change substantially. The ESD-t model is clearly preferred according to both the likelihood and information criteria, and it also delivers the best out-of-sample performance. While the SD-t model is preferred over PD-$\mathcal{N}$, highlighting the importance of capturing tails information, it is never favored over the ESD-t by in-sample criteria and performs worse in terms of joint density forecasting. This outcome is expected, as SD models omit the conditional covariance dynamics that capture an important part of the dependence structure. Nonetheless, all models exhibit broadly similar performance in terms of the factor and data MSEs, highlighting that all models deliver competitive point forecasts.

	PD-$\mathcal{N}$	SD-$\mathcal{N}$	ESD-$\mathcal{N}$	SD-t	ESD-t
DGP 1: $\mathcal{N}$, $\boldsymbol{P} = \mathbf{0}$					
LSR	**4.6529**	4.8268	4.6539	4.8248	4.6543
MSE(f_t)	**0.3325**	0.3379	0.3328	0.3384	0.3328
MSE($y_{1,t}$)	**0.5325**	0.5377	0.5328	0.5382	0.5328
DGP 2: t, $\boldsymbol{P} = \mathbf{0}$					
LSR	6.6715	6.7520	6.6734	6.0793	**5.8986**
MSE(f_t)	0.5933	0.5804	0.5900	0.5837	**0.5528**
MSE($y_{1,t}$)	1.1989	1.1878	1.1948	1.1919	**1.1622**
DGP 3: $\mathcal{N}$, $\boldsymbol{P} \neq \mathbf{0}$					
LSR	**4.6796**	4.8793	4.6810	4.8774	4.6814
MSE(f_t)	**0.3307**	0.3373	0.3315	0.3380	0.3315
MSE($y_{1,t}$)	**0.5493**	0.5538	0.5498	0.5545	0.5499
DGP 4: t, $\boldsymbol{P} \neq \mathbf{0}$					
LSR	6.7269	6.8219	6.7287	6.1639	**5.9543**
MSE(f_t)	0.5723	0.6164	0.5766	0.5924	**0.5491**
MSE($y_{1,t}$)	1.2233	1.2223	1.2232	1.2214	**1.1931**

Table 4: **Out-of-sample Monte Carlo results.**

5 Empirical application

5.1 Constructing Philadelphia Fed's composite indices

5.1.1 Coincident Economic Index

The Federal Reserve Bank of Philadelphia produces coincident indices for 50 US states as well as a national-level coincident economic index, which we refer to as the 'PHI CEI' in what follows. The original construction of these indices uses the DFM based on the Kalman filter in the spirit of Stock and Watson (1989) (see Crone and Clayton-Matthews, 2005, for details). As described in the introduction, in the construction of these indices the Philadelphia Fed treats all 2020 observations as missing, in order to avoid the excessive impact of the extreme observations during the COVID-19 pandemic. Here, we demonstrate how the ESD framework proposed in this paper can be used to construct a coincident index in a fully model-based and data-driven manner, without relying on any ad hoc exclusions of specific time periods.

In our analysis, we use the same constituent series as those employed in constructing the national-level PHI CEI, with data retrieved from the FRED database maintained by the St Louis Fed. Specifically, the coincident index is based upon nonfarm payroll employment, the unemployment rate, average weekly hours worked in manufacturing by production workers, and real personal income (defined as the sum of wages and salaries with proprietors' income deflated by the consumer price index). We work with monthly data for the sample period from January 1959 until August 2025. To ensure stationarity, we take log-differences of all series except the unemployment rate, for which we take first differences. Furthermore, we standardize all series to have zero mean and unit variance, as is typically done in the DFM literature.

First, we compare the fit across different model classes: (extended) score-driven with Gaussian and Student's t innovations and parameter-driven models. In addition, while we include a single factor ($r = 1$) for all specifications, we consider a range of DFM variants: stylized models without additional lags, as discussed in Section 2, and specifications that include m lags of the factor (with $m = 1, 2$)[6] and AR(p) dynamics for the idiosyncratic components (with $p = 1, 2$), where we set $m = p$ throughout. The estimation results are summarized in terms of likelihoods and information criteria, as reported in Table 5.

Three findings stand out from Table 5. First, (E)SD models with t-distributed in-

[6]To ensure identification, we restrict the factor loadings of the first series (employment), such that the common factor enters its measurement equation only contemporaneously. Such an identification constraint is standard in the DFM literature (see, e.g., Stock and Watson, 1989) and is also employed in the Philadelphia Fed methodology (Crone and Clayton-Matthews, 2005).

		PD-$\mathcal{N}$	SD-$\mathcal{N}$	ESD-$\mathcal{N}$	SD-t	ESD-t
$p = 0, m = 0$	Loglik	-3513.13	-4463.91	-3510.00	-2146.64	**-2064.73**
	BIC	7086.27	8987.82	7086.68	4359.95	**4202.79**
	cAIC	7095.27	8996.82	7096.68	4369.95	**4213.79**
$p = 1, m = 1$	Loglik	-3408.59	-4413.52	-3375.04	-2167.15	**-1942.87**
	BIC	6923.86	8933.70	6863.43	4447.65	**4005.75**
	cAIC	6939.86	8949.70	6880.43	4464.65	**4023.75**
$p = 2, m = 2$	Loglik	-3384.85	-4383.39	-3336.05	-2077.47	**-1953.81**
	BIC	6923.03	8920.12	6832.11	4314.94	**4074.30**
	cAIC	6946.03	8943.12	6856.11	4338.94	**4099.30**

Table 5: **Model fit comparison for PHI CEI constituents.** We report the log-likelihood value (Loglik) and Bayesian and consistent Akaike information criteria (BIC and cAIC) for parameter-driven (PD), score-driven (SD), and extended score-driven (ESD) models with either Gaussian ($\mathcal{N}$) or Student's t (t) distributions. m indicates the included number of lags of the common factors, whereas p is the autoregressive order assumed for the idiosyncratic components. All model specifications are without conditional heteroskedasticity and with $q = 0$.

novations outperform their Gaussian counterparts, with a very substantial difference in log-likelihoods and information criteria. Second, for both innovation distributions, the ESD model achieves a superior fit compared to the standard SD model, where again the difference is substantial especially in the Gaussian case. Third, the performance of the Gaussian ESD-DFM and PD-DFM is almost identical for the specification without lags ($p = 0, m = 0$). This is intuitive because in this case the ESD-$\mathcal{N}$ model nests the PD-$\mathcal{N}$ specification. However, when additional lags are considered, the ESD model clearly outperforms its PD counterpart.

Next, we undertake a more refined specification search for the ESD-t model. We explore a range of lag specifications with $p \in \{0, 1, 2, 3\}$ and $m \in \{0, 1, 2, 3\}$, considering models with and without conditional heteroskedasticity and distinguishing between robust and non-robust versions of the volatility update. The cAIC results are reported in Table 6.

We find that the specification with a single lag of the common factor in the measurement equation ($m = 1$) and AR(1) dynamics for the innovation terms ($p = 1$) is preferred, both for the homoskedastic and heteroskedastic model variants. Independent of the values of m and p, the model specifications with conditional heteroskedasticity are preferred over the homoskedastic model, with the 'GARCH'-type specification favored over the robust version. This latter result is particularly insightful, as it supports the interpretation of the COVID-19 period as representing "shifts to the distribution of existing shocks" (Ng, 2021), which is captured in our model through time-varying volatility; see also Figure 2 for the plot of the estimated volatility. Additionally, the preference for the non-robust volatility update further reinforces the idea that these shocks are better absorbed through volatility

	$p=0$	$p=1$	$p=2$	$p=3$
No conditional heteroskedasticity				
$m=0$	4213.79	4192.99	4187.82	4154.76
$m=1$	4125.90	**4023.75**	4107.27	4073.95
$m=2$	4132.22	4034.04	4099.30	4081.13
$m=3$	4141.21	4130.71	4119.81	4086.08
GARCH				
$m=0$	4074.27	4052.50	4048.23	4016.19
$m=1$	3969.16	**3866.32**	3874.51	3969.77
$m=2$	3970.52	3881.05	3879.46	3905.33
$m=3$	3968.61	3895.21	3885.98	3912.64
robust GARCH				
$m=0$	4137.15	4046.68	4109.88	4084.09
$m=1$	4042.79	**3946.63**	4046.10	4016.45
$m=2$	4046.08	3961.46	4043.78	4026.70
$m=3$	4050.36	3973.59	4063.34	4037.68

Table 6: **ESD-t model selection for the PHI CEI constituents.** We report the values of cAIC for different model specifications. The models are estimated using the full sample.

adjustments.

Figure 2 shows the predicted factor together with 95% confidence bands, as obtained from the preferred ESD-t model specification with $m=1$, $p=1$ and non-robust GARCH volatility dynamics. Periods with negative factor values align closely with the "official" US recessions based upon the turning points determined by the NBER business cycle dating committee. Indeed, the confidence bands indicate that throughout all recessions, the factor is statistically different from zero. By contrast, standard score- or observation-driven models cannot account for state uncertainty and typically only parameter estimation uncertainty is reported (Blasques et al., 2016).

Finally, we reconstruct the composite index in levels by integrating the estimated factor obtained from the ESD-t model (through reverse-differencing and exponentiation) following the approach of Crone and Clayton-Matthews (2005). Specifically, we apply this transformation to $\mu + \phi f_t$, where the coefficients μ and ϕ are chosen to align the mean and scale of the transformed factor with those of the log-differenced PHI CEI.[7] Following the Philadelphia Fed convention, we set June 2007 as base period with the index value normalized to 100. The reconstructed index is displayed alongside the official PHI CEI in Figure 3. Overall, the estimated factor closely resembles the PHI CEI. The main differences arise

[7]The coefficients μ and ϕ are obtained from a regression of the log-differenced PHI CEI on the estimated factor f_t, excluding six months during the COVID-19 period to mitigate the influence of outliers.

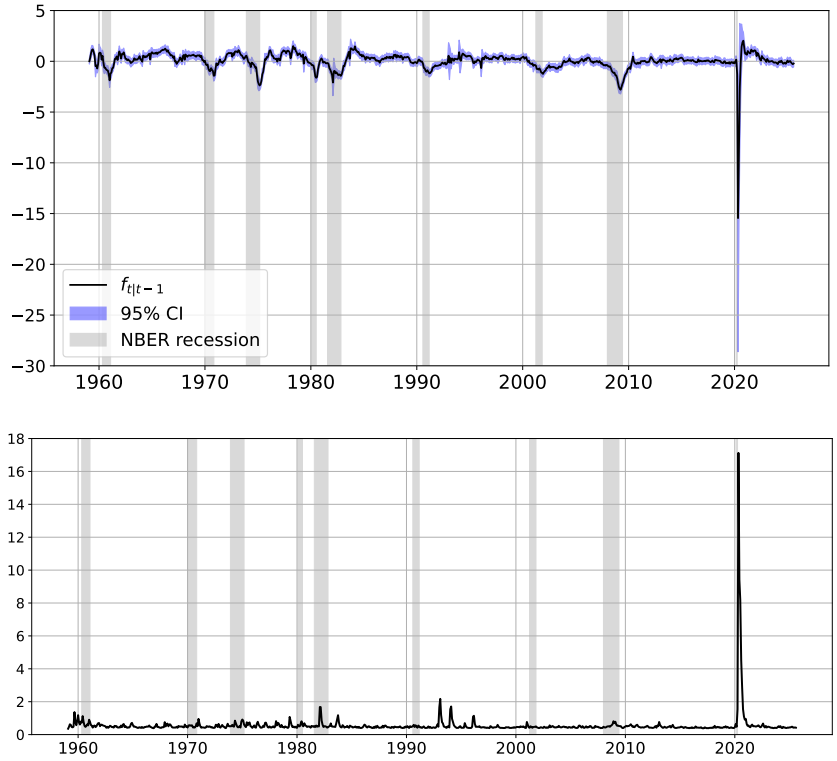


Figure 2: **Predicted factor together with the confidence intervals (top) and volatility (bottom) for the PHI CEI constituents.**

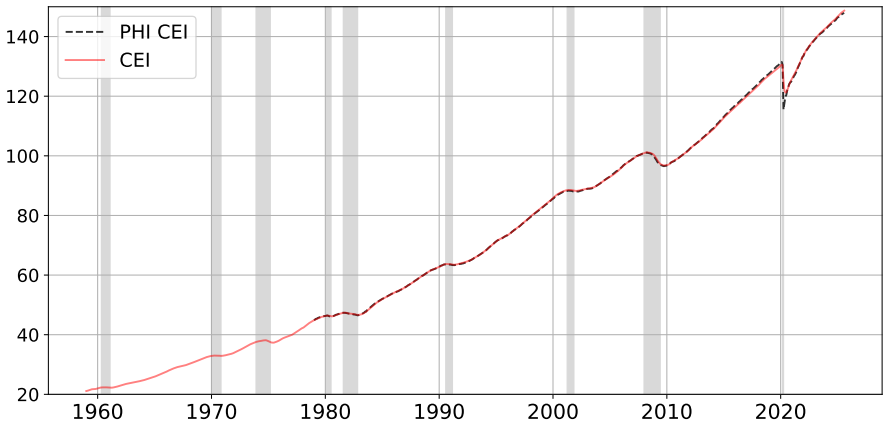


Figure 3: **Estimated CEI together with the PHI CEI.**

during periods of extreme observations. In particular, the estimated factor from the ESD-t model exhibits a smaller decline in 2020, which can be attributed to the robustness features in our model. Importantly, unlike the Philadelphia Fed's adjusted methodology, we do not treat the 2020 observations as missing when estimating the ESD-DFM.

5.1.2 Leading Economic Index

Until 2020, the Philadelphia Fed also produced both state- and national-level leading indices. However, because its methodology could not accommodate the extreme impact of the COVID-19 pandemic on initial unemployment claims, the Philadelphia Fed ultimately decided to discontinue the publication of these indices. In this subsection, we show that the ESD-DFM methodology effectively resolves this issue, ensuring the continued production of the leading index.

According to the Philadelphia Fed's methodology, the national leading economic index (PHI LEI) is designed to predict the six-month growth rate of the corresponding coincident index. In addition to the PHI CEI, the methodology utilizes four leading variables: state-level housing permits, state initial unemployment insurance claims, delivery times from the Institute for Supply Management manufacturing survey, and the interest rate spread. In the original methodology, a VAR model with these five variables is used to construct the leading index.

Instead of employing a VAR model, we estimate an ESD-DFM for the leading variables to mitigate the influence of extreme observations, and then use the extracted filtered factor to forecast the six-month growth rate of the PHI CEI. Specifically, we estimate

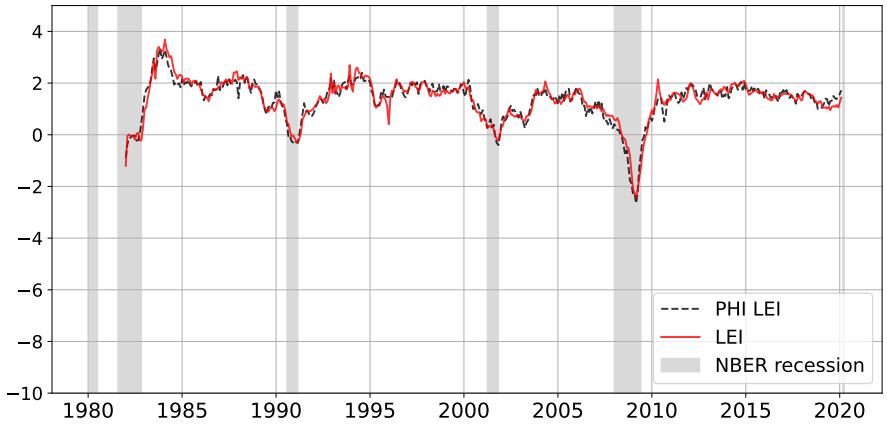
$$\log \frac{CEI_{t+6}}{CEI_t} = \beta_0 + \beta_c(L) \log \frac{CEI_t}{CEI_{t-1}} + \beta_l(L) f_t + \varepsilon_t, \tag{14}$$

where CEI_t is our estimated CEI as described in Section 5.1.1 and f_t is the estimated filtered factor obtained from the ESD-t specification applied to the set of leading variables. The number of lags is chosen using the BIC. Unlike the VAR, this approach imposes a factor structure on the leading variables, which is possibly related to the (leads of the) factor driving the coincident variables. The resulting predicted six-month growth rate constitutes our leading index. A similar approach to construct leading indices is used by Stock and Watson (2002a,b) in the context of high-dimensional factor models.[8]

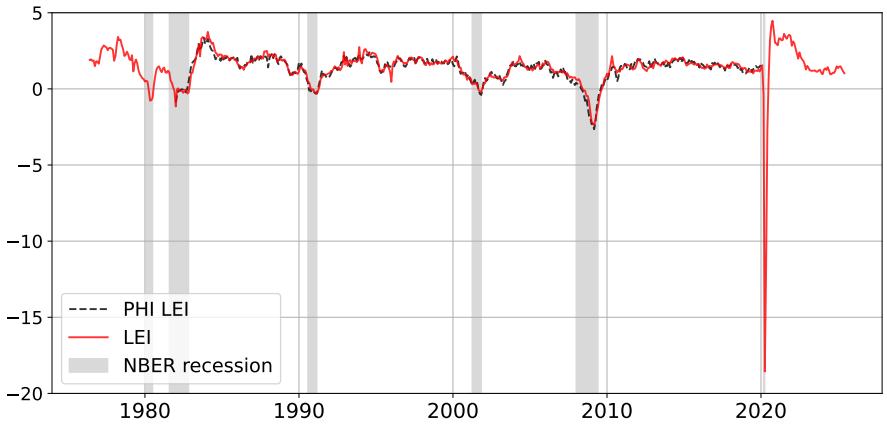
To compare the factor-based approach with the VAR-based method employed by the Philadelphia Fed, we first estimate our model using the sample period for which the PHI LEI is available (from January 1982 until February 2020) and construct the leading index using equation (14). For this purpose, we re-estimate the CEI using data up to February 2020 and estimate the ESD-t model for the leading variables over the same period. Model selection is performed using the cAIC, similar to the analysis in Section 5.1.1.

Figure 4a displays the estimated LEI using the factor-based approach along with the PHI LEI reported by the Philadelphia Fed. We observe that the estimated LEI closely

[8]See also Marcellino (2006) for a review of methods for constructing leading indices.



(a) Subsample from January 1982 until February 2020



(b) Full sample

Figure 4: **Estimated LEI together with the PHI LEI.**

tracks the official PHI LEI during the complete sample period. This indicates that the factor-augmented regression given by equation (14) can indeed be used as an alternative to the VAR approach for constructing leading indices.

Furthermore, since our ESD-DFM-t is robust to extreme observations, such as during the COVID-19 pandemic, unlike the Philadelphia FED we can continue to produce the LEI up to present times. For this purpose, we re-estimate the ESD-t specification on the full available sample period of the leading variables (from June 1976 until June 2025) and use the full-sample estimation of the CEI as discussed in Section 5.1.1. To produce the LEI, we use the regression estimates from the subsample analysis. Figure 4b displays the resulting LEI. We observe that the estimated LEI experiences a huge drop due to the extreme data observations at the beginning of the pandemic outbreak. However, this impact disappears rather quickly, with the LEI taking positive values again already in June 2020.

5.2 Constructing TCB composite indices

The Conference Board (TCB) also produces nationwide coincident and leading indices to signal peaks and troughs in the business cycle for major economies around the world, including the United States. These indices are widely followed and often attract substantial media attention. Unlike the Philadelphia Fed's approach, TCB methodology is model-free, and the indices are constructed as simple weighted averages of their constituent series. In this section, we demonstrate that although TCB methodology is straightforward to implement, their indices may not always be reliable. Specifically, we show that once the dynamic properties of the series are taken into account, the empirical puzzle of the diverging TCB coincident and leading indices is resolved, resulting in economically meaningful indices.

5.2.1 TCB Coincident Economic Index

TCB coincident economic index (TCB CEI) is constructed using four variables: nonfarm payroll employment, personal income excluding transfer payments, manufacturing and trade sales, and the industrial production index. The first two variables are the same as for the PHI CEI. We retrieve monthly observations from January 1959 until June 2025 and apply the necessary (log first-difference) transformations to estimate the DFM. In this section, we only briefly outline the results for TCB CEI, as the overall findings and methodology are similar to those discussed in Section 5.1.1. Detailed results on model fit comparison and model selection are presented in Appendix B. Overall, the ESD-DFM with Student's t distribution, $p = 2$ and $m = 1$, and GARCH-type conditional heteroskedasticity is selected.

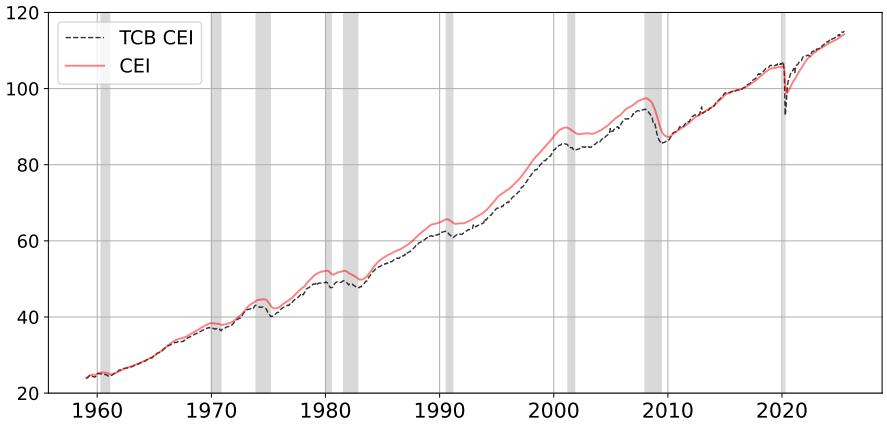


Figure 5: **Estimated CEI together with TCB CEI.**

Given the filtered factor, we reconstruct the CEI following the same approach applied for the PHI CEI in the previous section. Specifically, we integrate $\mu + \phi f_t$, where μ and

ϕ are the regression coefficients of the log-differenced TCB CEI on the filtered factor f_t. Following TCB, we set 2016 as the base year for the index level. The reconstructed CEI is displayed along with the official TCB CEI in Figure 5. The indices exhibit similar dynamics, albeit there are more discrepancies than in the case of the Philadelphia Fed. The estimated index appears somewhat smoother, likely due to accounting for dynamic features, and exhibits a smaller dip in 2020, which can be attributed to the robustness features in the model.

5.2.2 TCB Leading Economic Index

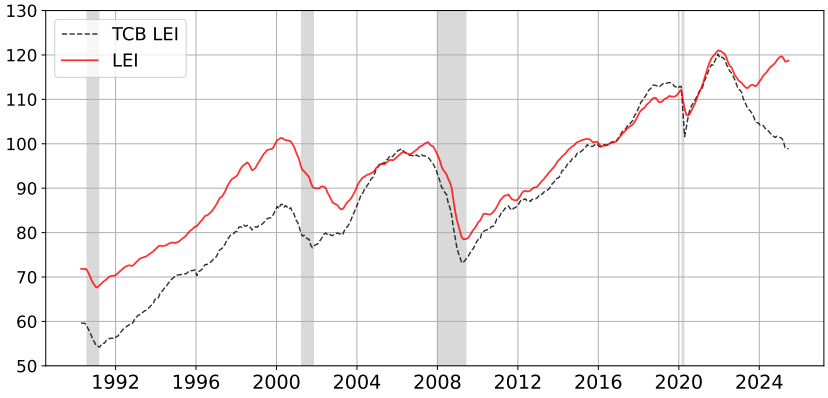
Similarly to the CEI, we construct TCB Leading Economic Index (TCB LEI), using the same constituent series and data transformations as those employed by TCB.[9] Specifically, the index consists of the average weekly hours worked by production workers in manufacturing industries (HKIM.O), average weekly initial claims for unemployment insurance (UNINSCE), manufacturers' new orders, consumer goods and materials (CNORCGD), ISM new order index (NAPMNO), manufacturers' new orders for non-defense capital goods excluding aircraft (NOEXCHD), building permits for new private housing units (HOUSATE), the S&P 500 index (500STK), leading credit index (BCILCIQ), interest rate spread (YST-NFF), and average consumer expectations for business and economic conditions (AVG-EXPQ). To maintain a balanced panel, we focus on the period from May 1990 to June 2025.

Among several alternatives, we select the same model specification as for the CEI in the previous section – an ESD-t model with $p = 2$, $m = 1$ and GARCH effects (see Table B.4 in Appendix). Figure 6a presents the estimated index in levels alongside The Conference Board's official LEI. Our constructed index and TCB LEI differ substantially at the end of the sample. Specifically, after August 2023 TCB LEI declines whereas our model-based index continues to rise.

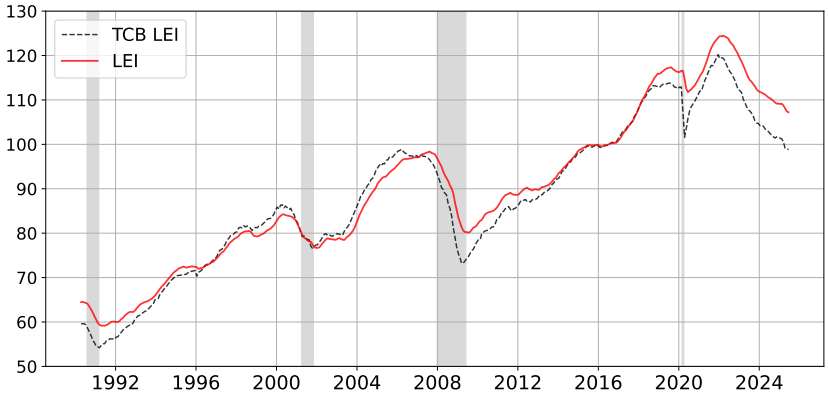
This discrepancy between our model-based index and TCB LEI sheds light on a broader puzzle that has recently emerged. Historically, TCB LEI and CEI have tended to move together, reflecting their role as indicators of the (expected) business cycle. However, in recent years this relationship has broken down: TCB CEI has recovered, while TCB LEI has continued to trend downward[10] (see Figure 1). Our fully model-based approach, which accounts for all relevant dynamic features, indicates an upward trajectory for the LEI once these effects are properly incorporated.

[9] `https://www.conference-board.org/data/bci/index.cfm?id=2160#BCI106`, `https://www.conference-board.org/data/bci/index.cfm?id=2154`

[10] This divergence phenomena has also received some media attention: Optima Capital Management (2023); State Street Global Advisors (2025)



(a) ESD-t with $p = 2, m = 1$



(b) ESD-$\mathcal{N}$ with $p = 0, m = 0$

Figure 6: **Estimated LEI together with TCB LEI.**

To better understand these differences, we also examine the index implied by an ESD-$\mathcal{N}$ model without lags ($p = 0$, $m = 0$), which is the closest to TCB methodology because it excludes both robustness features and dynamic propagation effects. The resulting index, shown in Figure 6b, closely resembles TCB LEI. Analysis based on additional, unreported results indicates that the primary driver of the divergence is the introduction of autoregressive effects in the innovations: once these dynamic effects are properly modeled, the index begins increasing after August 2023, whereas omitting them produces the downward trend observed in TCB LEI. Intuitively, failing to account for dynamics in the innovations introduces biases in the estimated loadings and, consequently, alters the interpretation of the underlying factor.

To gain more insight into the differences in factor dynamics, we examine the estimated loadings (Figure 7). For all series, both models assign the same sign to the loadings, but the magnitude of these contributions differs substantially. In particular, the ESD-$\mathcal{N}$ model without lags gives more weight to the ISM new orders index, interest rate spread, and

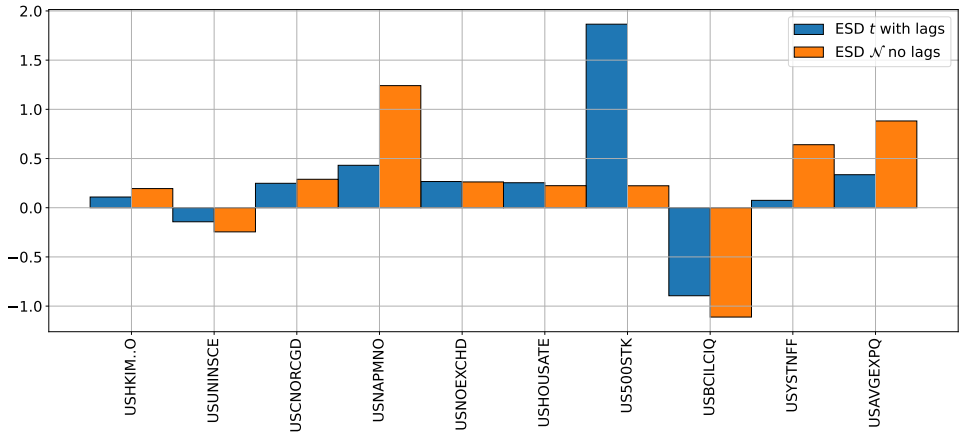


Figure 7: **Estimated loadings for LEI**

consumer expectations. By contrast, the ESD-t model with lags assigns a much larger weight to the S&P 500 index while almost no weight to the interest rate spread. The latter is somewhat surprising, as previous research has documented that the interest rate spread contains useful information for signaling future recessions (Harvey, 1988; Rudebusch and Williams, 2009). However, according to recent reporting,[11] it is questionable whether the interest rate spread is a reliable recession indicator this time. It is notable that according to our model, the interest rate spread co-moves less with other leading variables, indicating marginal relevance for summarizing the expected state of the economy. Overall, the factors' stronger reliance on the financial sector and lesser weight on manufacturing indicators contributes to the upward dynamics of the constructed LEI.

We conclude this section by investigating the predictive performance of TCB LEI and our model-based analogue. According to TCB, the CEI reflects current economic conditions, whereas the LEI is intended as a predictive tool that anticipates turning points in the business cycle. This interpretation of the LEI is similar in spirit to the definition of the LEI produced by the Philadelphia FED, where the PHI LEI is defined as a predictor of the six-month CEI growth rate. Therefore, we consider the predictive regression of the form given by equation (14), where we regress the six-month growth rate of TCB CEI on the monthly growth rates of the official TCB CEI and LEI (and their lags chosen by the BIC), instead of the filtered factor on the right-hand side of the equation. We then compare the results of this regression to those obtained when TCB LEI is replaced by our estimated LEI.

In Table 7, we report the in- and out-of-sample RMSEs for the fitted regressions, with the latter computed over different periods. To mitigate the impact of the COVID-19

[11] `https://people.duke.edu/~charvey/Media/2023/F_January_11_2023.pdf`
`https://people.duke.edu/~charvey/Media/2023/M_February_8_2023.pdf`

	In-Sample	Out-of-Sample			
		After 2021	After 2022	After 2023	After 2024
RMSE with TCB LEI	0.87	1.12	1.03	1.05	0.77
RMSE with our LEI	0.87	0.83	0.81	0.72	0.49
Ratio	0.99	0.75	0.79	0.68	0.64

Table 7: **Predictive regression result.** We report the RMSE×100 for the regression of TCB CEI six-month growth rate on the monthly percentage change of TCB CEI and of (i) TCB LEI and (ii) estimated LEI.

observations on the regression results and to assess out-of-sample performance, we estimate the regression using the data until December 2019. The results indicate that, in-sample, the difference in fit between the two regressions is marginal. However, out-of-sample, the regression using our estimated LEI yields RMSE reductions of 20% – 35% compared to the regression based on the official TCB LEI. These findings further highlight that the model-based construction of the LEI, which accounts for all relevant dynamic features, contains more predictive information (at least for the post-pandemic period) about TCB CEI than the official TCB LEI. Therefore, the proposed ESD-DFM allows us to construct the index that better 'leads' TCB CEI and, at the same time, resolves the divergent pattern between the CEI and LEI depicted in Figure 1.

6 Conclusion

We have introduced a generalized version of the score-driven dynamic factor model that nests parameter-driven Gaussian DFM, thereby bridging the gap between these two classes of models. Due to observation-driven specification, the model naturally accommodates features such as non-Gaussian innovations, nonlinear factor dynamics, and time-varying volatility, all within a frequentist framework. Our approach facilitates the construction of coincident and leading economic indices, such as those historically produced by the Federal Reserve Bank of Philadelphia, some of which were later discontinued. Applied to reconstruct The Conference Board's Coincident and Leading Economic Indices, the model demonstrates superior performance relative to both the Kalman filter–based state-space DFMs and standard score-driven DFMs. Importantly, the CEI and LEI estimated using our approach co-move throughout the entire sample, in contrast to the official indicators reported by The Conference Board. These results highlight the advantages of model-based indicator construction: by explicitly accounting for dynamic relationships and accounting for robustness, the approach can reveal the relative importance of different series and provide economically meaningful composite indicators.

Appendices

A Additional results for the illustrative model

A.1 Observational equivalence

In this appendix, we show the observational equivalence between the single-factor PD-DFM and ESD-DFM under Gaussian assumption considered in Section 2. For that, we will show that the two models yield the same predictive distribution for the data $\boldsymbol{y}_{t+1}$ conditional on information $\mathcal{F}_t$.

Recall, that for the PD-DFM given by equations (1) and (2), the Kalman filter recursion after reaching steady-state is given by

$$\widehat{f}_{t+1|t} = \phi \widehat{f}_{t|t-1} + \boldsymbol{k}(\boldsymbol{y}_t - \boldsymbol{\lambda}\widehat{f}_{t|t-1}) \text{ with } \boldsymbol{k} = \frac{\phi p}{1+pN}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1}, \tag{15}$$

where $\widehat{f}_{t+1|t} = \mathbb{E}[f_{t+1}|\mathcal{F}_t]$. This yields the moments of the data given by

$$\begin{aligned}
\mathbb{E}[\boldsymbol{y}_{t+1}|\mathcal{F}_t] &= \boldsymbol{\lambda}\mathbb{E}[f_{t+1}|\mathcal{F}_t] = \boldsymbol{\lambda}\widehat{f}_{t+1|t}, \\
\mathbb{V}[\boldsymbol{y}_{t+1}|\mathcal{F}_t] &= \mathbb{V}[f_{t+1}|\mathcal{F}_t]\boldsymbol{\lambda}\boldsymbol{\lambda}^\top + \boldsymbol{\Sigma} = p\boldsymbol{\lambda}\boldsymbol{\lambda}^\top + \boldsymbol{\Sigma}.
\end{aligned}$$

Let us now derive the conditional mean and the variance of the single-factor ESD-DFM. Recall that the factor updating equation in the ESD-DFM in the Gaussian case is given by

$$f_{t+1} = bf_t + a\frac{1}{N}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t - \boldsymbol{\lambda}f_t) + c\frac{1}{N}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_{t+1} - \boldsymbol{\lambda}f_{t+1}). \tag{16}$$

Thus, the conditional mean and the variance of the factor are

$$\begin{aligned}
\mathbb{E}[f_{t+1}|\mathcal{F}_t] &= bf_t + a\frac{1}{N}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t - \boldsymbol{\lambda}f_t) = f_{t+1|t}, \\
\mathbb{V}[f_{t+1}|\mathcal{F}_t] &= \mathbb{E}\left[\frac{c^2}{N^2}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_{t+1}\boldsymbol{\varepsilon}_{t+1}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda}\middle|\mathcal{F}_t\right] = \frac{c^2}{N^2}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda} = \frac{c^2}{N}.
\end{aligned}$$

Here and throughout, we avoid the 'hat'-notation in the predicted factor $f_{t+1|t}$ to distinguish it from the predicted factors obtained from the Kalman filter under the PD-DFM framework. Given the conditional moments of the factors, we can derive the conditional expectation of the data:

$$\mathbb{E}[\boldsymbol{y}_{t+1}|\mathcal{F}_t] = \boldsymbol{\lambda}\mathbb{E}[f_{t+1}|\mathcal{F}_t] = \boldsymbol{\lambda}f_{t+1|t},$$

and the conditional variance of the data:

$$\begin{aligned}
\mathbb{V}[\boldsymbol{y}_{t+1}|\mathcal{F}_t] &= \mathbb{E}[(\boldsymbol{y}_{t+1} - \mathbb{E}[\boldsymbol{y}_{t+1}|\mathcal{F}_t])\,(\boldsymbol{y}_{t+1} - \mathbb{E}[\boldsymbol{y}_{t+1}|\mathcal{F}_t])^\top\,|\mathcal{F}_t] \\
&= \mathbb{E}[(\boldsymbol{\lambda}(f_{t+1} - f_{t+1|t}) + \boldsymbol{\varepsilon}_{t+1})(\boldsymbol{\lambda}(f_{t+1} - f_{t+1|t}) + \boldsymbol{\varepsilon}_{t+1})^\top|\mathcal{F}_t] \\
&= \boldsymbol{\lambda}\mathbb{E}[(f_{t+1} - f_{t+1|t})(f_{t+1} - f_{t+1|t})^\top|\mathcal{F}_t]\boldsymbol{\lambda}^\top + \boldsymbol{\Sigma} + 2\mathbb{E}[\boldsymbol{\lambda}(f_{t+1} - f_{t+1|t})\boldsymbol{\varepsilon}_{t+1}^\top|\mathcal{F}_t] \\
&= \boldsymbol{\lambda}\mathbb{V}[f_{t+1}|\mathcal{F}_t]\boldsymbol{\lambda}^\top + \boldsymbol{\Sigma} + 2c\frac{1}{N}\mathbb{E}[\boldsymbol{\lambda}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_{t+1}\boldsymbol{\varepsilon}_{t+1}^\top|\mathcal{F}_t] \\
&= \frac{c^2}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top + \boldsymbol{\Sigma} + 2c\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top \\
&= \frac{c^2 + 2c}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top + \boldsymbol{\Sigma}.
\end{aligned}$$

Therefore, comparing the conditional moments for the two classes of models, the Gaussian PD-DFM and ESD-DFM are observationally equivalent if

$$p = \frac{c^2 + 2c}{N} \text{ and } \widehat{f}_{t+1} = f_{t+1|t}.$$

To show this, it is convenient to rewrite the prediction step of the factor in the ESD-DFM in the form of recursion:

$$\begin{aligned}
f_{t+1|t} &= a\frac{1}{N}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{y}_t + (b-a)f_t \\
&= a\frac{1}{N}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{y}_t + \frac{b-a}{1+c}\left(f_{t|t-1} + c\frac{1}{N}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{y}_t\right) \\
&= \left(a + \frac{c(b-a)}{1+c}\right)\frac{1}{N}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{y}_t + \frac{b-a}{1+c}f_{t|t-1} \\
&= \left(a + \frac{c(b-a)}{1+c}\right)\frac{1}{N}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{y}_t - \boldsymbol{\lambda}f_{t|t-1}\right) + \left[\frac{b-a}{1+c} + a + \frac{c(b-a)}{1+c}\right]f_{t|t-1} \\
&= \underbrace{\frac{a+cb}{1+c}}_{=:\widetilde{a}}\frac{1}{N}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{y}_t - \boldsymbol{\lambda}f_{t|t-1}\right) + \frac{b-a+a(1+c)+c(b-a)}{1+c}f_{t|t-1} \\
&= \widetilde{a}\frac{1}{N}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{y}_t - \boldsymbol{\lambda}f_{t|t-1}\right) + bf_{t|t-1}, \qquad (17)
\end{aligned}$$

where we used the identification restriction $\frac{1}{N}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda} = 1$ for the first equality and the updating step given by equation (6) in the second line.

Now, comparing the steady-state Kalman filter recursion given by equation (15) with equation (17), we can use the induction argument, from which it follows that $\widehat{f}_{t+1|t} = f_{t+1|t}$ if $b = \phi$ and

$$\widetilde{a} = \frac{\phi pN}{1 + pN}.$$

On the other hand, from equating the conditional variances, we have that $pN = c^2 + 2c$. Substituting this expression, $b = \phi$, and the definition of $\widetilde{a}$ into the equation above, we obtain

$$\frac{a + cb}{1 + c} = \frac{b(c^2 + 2c)}{1 + c^2 + 2c}.$$

Solving this equation for c, we obtain that when

$$c = \frac{a}{b - a}, \text{ with } b \neq a,$$

the ESD-DFM is observationally equivalent to the PD-DFM.

A.2 State-space induced restriction

In this appendix, we show that the PD-DFM is more restrictive than the ESD-DFM. For simplicity, we focus on the illustrative single-factor specification considered in Section 2. The argument parallels the well-known univariate result that the local level model and AR(1)+noise model correspond to restricted AR(I)MA processes with AR and MA coefficients of opposite signs. In this section, we extend this insight to the multivariate setting in the context of dynamic factor models.

Consider the PD-DFM given by equations (1) and (2). Applying the lag operator $(1 - \phi L)$ to the observation equation yields

$$\boldsymbol{y}_{t+1} - \phi \boldsymbol{y}_t = \boldsymbol{\lambda} f_{t+1} + \boldsymbol{\varepsilon}_{t+1} - \phi(\boldsymbol{\lambda} f_t + \boldsymbol{\varepsilon}_t) = \boldsymbol{\lambda}\eta_{t+1} + \boldsymbol{\varepsilon}_{t+1} - \phi\boldsymbol{\varepsilon}_t.$$

The error term $\boldsymbol{u}_{t+1} := \boldsymbol{\lambda}\eta_{t+1} + \boldsymbol{\varepsilon}_{t+1} - \phi\boldsymbol{\varepsilon}_t$ has a simple autocorrelation structure, i.e., since η_t and $\boldsymbol{\varepsilon}_t$ are independent, its first lag autocovariance is

$$\mathbb{C}ov(\boldsymbol{u}_{t+1}, \boldsymbol{u}_t) = \mathbb{C}ov(\boldsymbol{\lambda}\eta_{t+1} + \boldsymbol{\varepsilon}_{t+1} - \phi\boldsymbol{\varepsilon}_t, \boldsymbol{\lambda}\eta_t + \boldsymbol{\varepsilon}_t - \phi\boldsymbol{\varepsilon}_{t-1}) = -\phi\boldsymbol{\Sigma},$$

whereas for all lags $k > 1$ we obtain $\mathbb{C}ov(\boldsymbol{u}_{t+k}, \boldsymbol{u}_t) = 0$. Hence, $\boldsymbol{u}_t$ displays the autocorrelation pattern of a restricted VMA(1) process: the only nonzero autocovariance beyond lag zero occurs at lag one and is constrained to be proportional to $\boldsymbol{\Sigma}$ with a sign opposite to the autoregressive parameter ϕ. This is analogous to the univariate AR(1) + noise model, which is equivalent to the restricted ARMA(1,1) process. This implies that the PD-DFM cannot generate VARMA processes with positive VAR(1) and VMA(1) coefficients.

At the same time, we can also rewrite the ESD-DFM as

$$\begin{aligned}\boldsymbol{y}_{t+1} - b\boldsymbol{y}_t &= a\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_t + c\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}_{t+1} + \boldsymbol{\varepsilon}_{t+1} - b\boldsymbol{\varepsilon}_t \\ &= \left(\boldsymbol{I} + c\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{\varepsilon}_{t+1} - b\left(\boldsymbol{I} - \frac{a}{b}\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{\varepsilon}_t \\ &=: \boldsymbol{u}_{t+1}.\end{aligned}$$

Clearly, in this example, the ESD-DFM corresponds to a restricted VARMA process, with the first lag autocovariance of the MA component given by

$$\begin{aligned}\mathbb{C}ov(\boldsymbol{u}_{t+1}, \boldsymbol{u}_t) &= -b\left(\boldsymbol{I} + c\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{\Sigma}\left(\boldsymbol{I} - \frac{a}{b}\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-1}\right)^\top \\ &= -b\left(\boldsymbol{\Sigma} + c\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top\right)\left(\boldsymbol{I} - \frac{a}{b}\frac{1}{N}\boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top\right) \\ &= -b\boldsymbol{\Sigma} + (a - bc + ac)\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^\top.\end{aligned}$$

Under the equivalence restriction $c = \frac{a}{b-a}$, the second term vanishes and the first lag autocovariance of the MA component coincides exactly with that of the PD-DFM. When $c \neq \frac{a}{b-a}$, the ESD-DFM generates a richer autocovariance structure, incorporating an off-diagonal rank-one component in addition to the diagonal part proportional to $\boldsymbol{\Sigma}$. Moreover, the presence of this additional term enables the ESD-DFM to generate second-order dynamics resembling VARMA processes with both positive VAR(1) and positive VMA(1) parameters, which is not possible within the PD-DFM framework.

B Additional empirical results

B.1 TCB CEI constituents

		PD-$\mathcal{N}$	SD-$\mathcal{N}$	ESD-$\mathcal{N}$	SD-t	ESD-t
$p=0, m=0$	Loglik	-3775.43	-4369.93	-3771.72	-2549.80	**-2397.76**
	BIC	7610.86	8799.84	7610.10	5166.25	**4868.84**
	CAIC	7619.86	8808.84	7620.10	5176.25	**4879.84**
$p=1, m=1$	Loglik	-3708.84	-4346.52	-3703.15	-2511.63	**-2296.73**
	BIC	7524.33	8799.70	7519.62	5136.58	**4713.45**
	CAIC	7540.33	8815.70	7536.62	5153.58	**4731.45**
$p=2, m=2$	Loglik	-3689.64	-4307.08	-3687.18	-2481.94	**-2238.89**
	BIC	7532.59	8767.47	7534.33	5123.85	**4644.42**
	CAIC	7555.59	8790.47	7558.33	5147.85	**4669.42**

Table B.1: **Model fit comparison for TCB CEI constituents.** We report the log-likelihood value (Loglik) and Bayesian and consistent Akaike information criteria (BIC and CAIC) for the parameter-driven (PD), score-driven (SD), and extended score-driven (ESD) models with either Gaussian ($\mathcal{N}$) or Student's t (t) distributions. All model specification are without conditional heteroskedasticity and with $q=0$.

	$p=0$	$p=1$	$p=2$	$p=3$
No conditional heteroskedasticity				
$m=0$	4879.84	4910.20	4763.49	4746.73
$m=1$	4770.99	4731.45	4680.74	4731.27
$m=2$	4769.46	4739.12	**4669.42**	4685.48
$m=3$	4781.85	4751.32	4684.13	4702.00
GARCH				
$m=0$	4748.90	4756.47	4620.24	4606.23
$m=1$	4618.40	4585.16	**4562.95**	4599.21
$m=2$	4623.83	4599.77	4572.77	4581.81
$m=3$	4637.74	4615.03	4590.25	4598.45
robust GARCH				
$m=0$	4802.53	4825.67	4685.74	4670.95
$m=1$	4685.42	4648.60	**4605.89**	4657.99
$m=2$	4689.54	4660.77	4609.54	4625.21
$m=3$	4705.06	4676.91	4627.15	4642.28

Table B.2: **ESD-t model selection for the TCB CEI constituents.** We report the values of CAIC for different model specifications. The models are estimated using the full sample.

			EMPNAGE	PILESTD	INPRODG	MANTRAD
a	0.132	$\lambda_{0,i}$	0.261	0.465	0.964	1.074
	(0.072)		(0.035)	(0.082)	(0.069)	(0.136)
b	0.942	$\lambda_{1,i}$	0.000	-0.199	-0.697	-0.886
	(0.014)		–	(0.056)	(0.062)	(0.092)
c	0.802	σ_i^2	0.083	0.769	0.531	1.000
	(0.071)		(0.015)	(0.094)	(0.068)	–
ν	4.889	$p_{1,i}$	0.049	0.122	0.043	-0.375
	(0.616)		(0.090)	(0.046)	(0.066)	(0.055)
ω	0.152	$p_{2,i}$	0.146	0.003	0.134	-0.207
	(0.047)		(0.052)	(0.038)	(0.071)	(0.052)
α	0.282					
	(0.078)					
β	0.334					
	(0.135)					

Table B.3: **Parameter estimates for the ESD-t model for TCB CEI constituents**

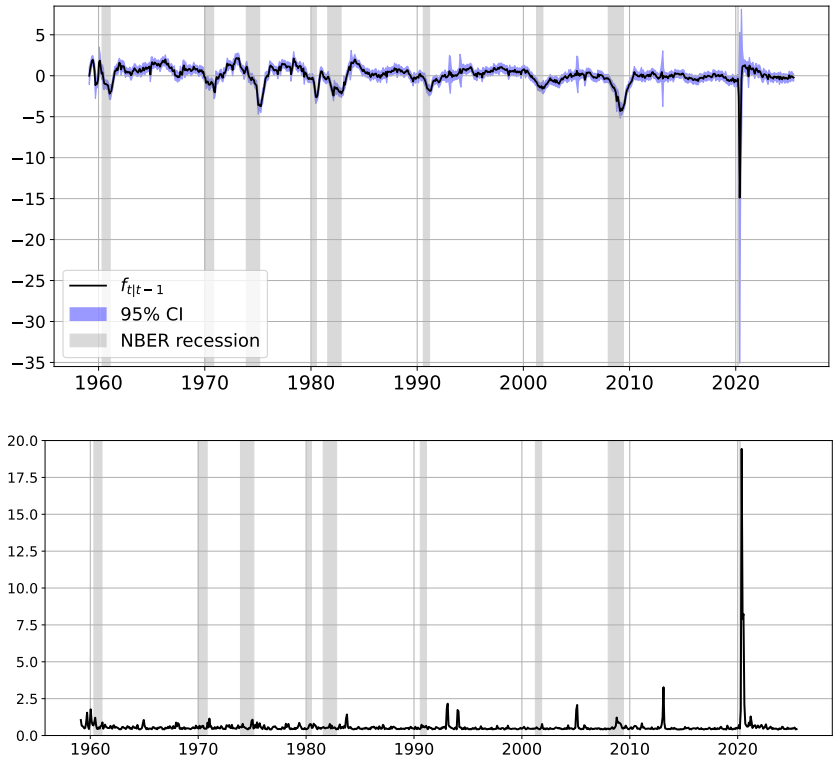


Figure B.1: **Predicted factor with the confidence interval (top) and volatility (bottom) for TCB CEI constituents.**

B.2 TCB LEI constituents

	$p=0$	$p=1$	$p=2$	$p=3$
No conditional heteroskedasticity				
$m=0$	8841.40	6806.68	6756.84	6801.58
$m=1$	9318.14	6759.13	**6710.77**	6749.15
$m=2$	9358.01	6805.02	6757.08	6789.92
$m=3$	9411.68	6857.00	6809.35	6844.47
GARCH				
$m=0$	8527.83	6688.90	6654.59	6701.46
$m=1$	9157.88	6635.98	**6598.56**	6636.98
$m=2$	8931.94	6685.18	6647.18	6681.66
$m=3$	9276.31	6738.56	6700.33	6736.84
robust GARCH				
$m=0$	9175.18	6740.09	6702.39	6749.96
$m=1$	9191.08	6689.24	**6649.09**	6688.47
$m=2$	9233.94	6737.88	6698.12	6734.00
$m=3$	8930.53	6791.89	6752.38	6789.99

Table B.4: **ESD-t model selection for the TCB LEI constituents.** We report the values of CAIC for different model specifications. The models are estimated using the full sample.

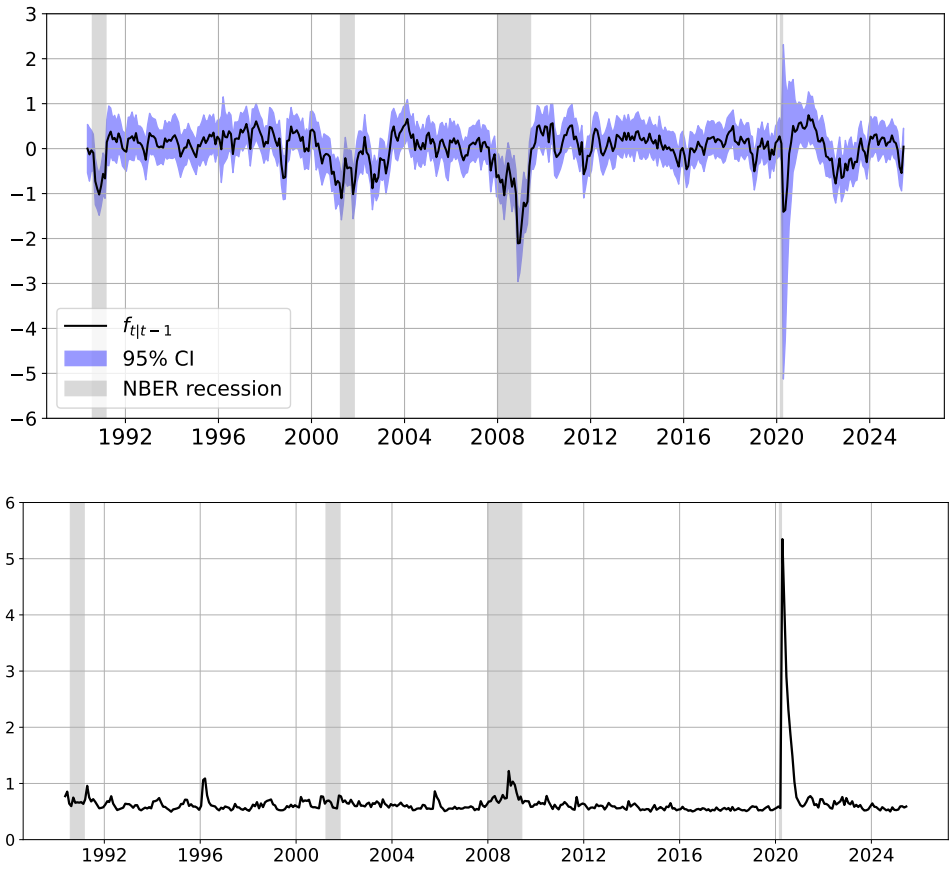


Figure B.2: **Predicted factor with the confidence interval (top) and volatility (bottom) for TCB LEI constituents.**

C Theoretical results

Proof of Lemma 1. Following Bai and Wang (2015, Proposition 1), we focus on the case $m = 1$. Overall, the structure of the proof is largely the same. The key distinction in our setting is that the factors are observation-driven, and under the extended specification the innovations enter the transition equation contemporaneously.

We first rewrite the dynamic factor in static form as $\boldsymbol{F}_t = (\boldsymbol{f}_t^\top, \boldsymbol{f}_{t-1}^\top)^\top$. Because the factor is identified only up to a rotation, we rotate the stacked state vector using an $((m+1)r \times (m+1)r)$ matrix. For $m = 1$, this yields

$$\begin{bmatrix} \boldsymbol{K} & \boldsymbol{L} \\ \boldsymbol{M} & \boldsymbol{N} \end{bmatrix} \begin{bmatrix} \boldsymbol{f}_t \\ \boldsymbol{f}_{t-1} \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{f}}_t \\ \tilde{\boldsymbol{f}}_{t-1} \end{bmatrix}.$$

It follows that for all t, it hols that $(\boldsymbol{K} - \boldsymbol{N})\boldsymbol{f}_{t-1} = \boldsymbol{M}\boldsymbol{f}_t - \boldsymbol{L}\boldsymbol{f}_{t-2}$. Recalling that $\boldsymbol{f}_{t+1} = \boldsymbol{B}(L)\boldsymbol{f}_t + \boldsymbol{A}\boldsymbol{s}_t^{\boldsymbol{f}} + \boldsymbol{C}\tilde{\boldsymbol{s}}_{t+1}^{\boldsymbol{f}}$, we have

$$(\boldsymbol{K} - \boldsymbol{N})(\boldsymbol{B}(L)\boldsymbol{f}_{t-2} + \boldsymbol{A}\boldsymbol{s}_{t-2}^{\boldsymbol{f}} + \boldsymbol{C}\tilde{\boldsymbol{s}}_{t-1}^{\boldsymbol{f}}) = \boldsymbol{M}(\boldsymbol{B}(L)\boldsymbol{f}_{t-1} + \boldsymbol{A}\boldsymbol{s}_{t-1}^{\boldsymbol{f}} + \boldsymbol{C}\tilde{\boldsymbol{s}}_t^{\boldsymbol{f}}) - \boldsymbol{L}\boldsymbol{f}_{t-2}.$$

Regrouping the terms and plugging in $\tilde{\boldsymbol{s}}_t^{\boldsymbol{f}} = \boldsymbol{S}_0\boldsymbol{\Lambda}_0^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{u}_t := \boldsymbol{Q}\boldsymbol{u}_t$, we obtain

$$\begin{aligned}
&((\boldsymbol{K} - \boldsymbol{N})\boldsymbol{B}(L) + \boldsymbol{L})\boldsymbol{f}_{t-2} + (\boldsymbol{K} - \boldsymbol{N})(\boldsymbol{A}\boldsymbol{s}_{t-2}^{\boldsymbol{f}} + \boldsymbol{C}\boldsymbol{Q}\boldsymbol{u}_{t-1}) \\
&\qquad = \boldsymbol{M}(\boldsymbol{B}(L)(\boldsymbol{A}\boldsymbol{s}_{t-2}^{\boldsymbol{f}} + \boldsymbol{B}(L)\boldsymbol{f}_{t-2} + \boldsymbol{C}\boldsymbol{Q}\boldsymbol{u}_{t-1}) + \boldsymbol{A}\boldsymbol{s}_{t-1}^{\boldsymbol{f}} + \boldsymbol{C}\boldsymbol{Q}\boldsymbol{u}_t). \\
&((\boldsymbol{K} - \boldsymbol{N})\boldsymbol{B}(L) - \boldsymbol{M}\boldsymbol{B}^2(L) + \boldsymbol{L})\boldsymbol{f}_{t-2} - \boldsymbol{M}\boldsymbol{A}\boldsymbol{s}_{t-1}^{\boldsymbol{f}} + ((\boldsymbol{K} - \boldsymbol{N}) - \boldsymbol{M}\boldsymbol{B}(L))\boldsymbol{A}\boldsymbol{s}_{t-2}^{\boldsymbol{f}} \\
&\qquad - \boldsymbol{M}\boldsymbol{C}\boldsymbol{Q}\boldsymbol{u}_t + (\boldsymbol{K} - \boldsymbol{N} - \boldsymbol{M})\boldsymbol{C}\boldsymbol{Q}\boldsymbol{u}_{t-1} = 0.
\end{aligned}$$

First, we consider the case $\boldsymbol{C} > 0$. We notice that $\boldsymbol{u}_t$ is uncorrelated with $\boldsymbol{u}_{t-1}$ and $((\boldsymbol{K} - \boldsymbol{N})\boldsymbol{B}(L) - \boldsymbol{M}\boldsymbol{B}^2(L) + \boldsymbol{L})\boldsymbol{f}_{t-2} - \boldsymbol{M}\boldsymbol{A}\boldsymbol{s}_{t-1}^{\boldsymbol{f}} + ((\boldsymbol{K} - \boldsymbol{N}) - \boldsymbol{M}\boldsymbol{B}(L))\boldsymbol{A}\boldsymbol{s}_{t-2}^{\boldsymbol{f}}$. Hence, $\boldsymbol{M}\boldsymbol{C}\boldsymbol{Q}\boldsymbol{u}_t = 0$ or $\boldsymbol{M} = 0$ because $\mathbb{V}ar(\boldsymbol{u}_t) > 0$, $\boldsymbol{C} > 0$ and $\boldsymbol{S}_0 \neq 0$, $\boldsymbol{\Lambda}_0 \neq 0$ and $\boldsymbol{\Sigma} < \infty$. Then, we have

$$((\boldsymbol{K} - \boldsymbol{N})\boldsymbol{B}(L) + \boldsymbol{L})\boldsymbol{f}_{t-2} + (\boldsymbol{K} - \boldsymbol{N})\boldsymbol{A}\boldsymbol{s}_{t-2}^{\boldsymbol{f}} + (\boldsymbol{K} - \boldsymbol{N})\boldsymbol{C}\boldsymbol{Q}\boldsymbol{u}_{t-1} = 0.$$

Similarly, $\boldsymbol{u}_{t-1}$ is uncorrelated with $((\boldsymbol{K} - \boldsymbol{N})\boldsymbol{B} + \boldsymbol{L})\boldsymbol{f}_{t-2}$ and with $(\boldsymbol{K} - \boldsymbol{N})\boldsymbol{A}\boldsymbol{s}_{t-2}^{f}$, which implies $\boldsymbol{K} = \boldsymbol{N}$. Consequently, $\boldsymbol{L}\boldsymbol{f}_{t-2} = 0$ for all $\boldsymbol{f}_{t-2}$, and therefore $\boldsymbol{L} = 0$. Hence, the model is identified up to a $q \times q$ rotation.

For $\boldsymbol{C} = 0$, we have

$$((\boldsymbol{K} - \boldsymbol{N})\boldsymbol{B}(L) - \boldsymbol{M}\boldsymbol{B}^2(L) + \boldsymbol{L})\boldsymbol{f}_{t-2} - \boldsymbol{M}\boldsymbol{A}\boldsymbol{s}_{t-1}^{\boldsymbol{f}} + ((\boldsymbol{K} - \boldsymbol{N}) - \boldsymbol{M}\boldsymbol{B}(L))\boldsymbol{A}\boldsymbol{s}_{t-2}^{\boldsymbol{f}} = 0,$$

and, using similar arguments, we obtain that $\boldsymbol{M} = 0$, $\boldsymbol{K} = \boldsymbol{N}$ and $\boldsymbol{L} = 0$, which finishes the proof. ■

Proof of Lemma 2. The conditional variance of the data is of the form

$$\begin{aligned}
\mathbb{V}ar(\boldsymbol{y}_t|\mathcal{F}_{t-1}) &= \mathbb{V}ar(\boldsymbol{\Lambda}_0\boldsymbol{f}_t + \boldsymbol{u}_t|\mathcal{F}_{t-1}) = \mathbb{V}ar(\boldsymbol{\Lambda}_0\boldsymbol{C}\boldsymbol{S}_0\boldsymbol{\Lambda}_0^\top\boldsymbol{\Sigma}^{-1}h_t\boldsymbol{u}_t + h_t\boldsymbol{u}_t|\mathcal{F}_{t-1}) && (18)\\
&= h_t^2\mathbb{V}ar((\boldsymbol{\Lambda}_0\boldsymbol{C}\boldsymbol{S}_0\boldsymbol{\Lambda}_0^\top\boldsymbol{\Sigma}^{-1} + \boldsymbol{I}_N)\boldsymbol{u}_t|\mathcal{F}_{t-1}) && (19)\\
&= h_t^2(\boldsymbol{\Lambda}_0\boldsymbol{C}\boldsymbol{S}_0\boldsymbol{\Lambda}_0^\top\boldsymbol{\Sigma}^{-1} + \boldsymbol{I}_N)\boldsymbol{\Sigma}(\boldsymbol{\Lambda}_0\boldsymbol{C}\boldsymbol{S}_0\boldsymbol{\Lambda}_0^\top\boldsymbol{\Sigma}^{-1} + \boldsymbol{I}_N)^\top && (20)\\
&= h_t^2(\boldsymbol{\Lambda}_0\boldsymbol{C}\boldsymbol{S}_0\boldsymbol{\Lambda}_0^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}_0\boldsymbol{S}_0\boldsymbol{C}\boldsymbol{\Lambda}_0^\top + 2\boldsymbol{\Lambda}_0\boldsymbol{C}\boldsymbol{S}_0\boldsymbol{\Lambda}_0^\top + \boldsymbol{\Sigma}), && (21)
\end{aligned}$$

where $\boldsymbol{S}_0 = (\boldsymbol{\Lambda}_0^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}_0)^{-1}$. Further assume that $h_t^2(\boldsymbol{\theta}) = h_t^2(\boldsymbol{\theta}_0)$ almost surely.

Assume that there exists another $\tilde{\boldsymbol{C}} = \text{diag}(\tilde{c}_1, \dots, c_r)$ with $c_j \neq \tilde{c}_j$ for all $j = 1, \dots r$ which leads to the same conditional variance. Then, given the identifiability of h_t^2, $\boldsymbol{\Lambda}_0$ and $\boldsymbol{\Sigma}$ and $h_t^2 > 0$, we have

$$\begin{aligned}
\boldsymbol{\Lambda}_0(\boldsymbol{C}\boldsymbol{S}_0\boldsymbol{C} + 2\boldsymbol{C}\boldsymbol{S}_0)\boldsymbol{\Lambda}_0^\top &= \boldsymbol{\Lambda}_0(\tilde{\boldsymbol{C}}\boldsymbol{S}_0\tilde{\boldsymbol{C}} + 2\tilde{\boldsymbol{C}}\boldsymbol{S}_0)\boldsymbol{\Lambda}_0^\top, && (22)\\
\boldsymbol{\Lambda}_0(\boldsymbol{C}\boldsymbol{S}_0\boldsymbol{C} - \tilde{\boldsymbol{C}}\boldsymbol{S}_0\tilde{\boldsymbol{C}})\boldsymbol{\Lambda}_0^\top + 2\boldsymbol{\Lambda}_0(\boldsymbol{C}\boldsymbol{S}_0 - \tilde{\boldsymbol{C}}\boldsymbol{S}_0)\boldsymbol{\Lambda}_0^\top &= 0. && (23)
\end{aligned}$$

Let us denote $\boldsymbol{M} = \boldsymbol{C}\boldsymbol{S}_0^{1/2}$ and $\tilde{\boldsymbol{M}} = \tilde{\boldsymbol{C}}\boldsymbol{S}_0^{1/2}$. Then, we have

$$\begin{aligned}
\boldsymbol{\Lambda}_0(\boldsymbol{M}^2 - \tilde{\boldsymbol{M}}^2)\boldsymbol{\Lambda}_0^\top + 2\boldsymbol{\Lambda}_0(\boldsymbol{M} - \tilde{\boldsymbol{M}})\boldsymbol{S}_0^{1/2}\boldsymbol{\Lambda}_0^\top &= 0, && (24)\\
\boldsymbol{\Lambda}_0(\boldsymbol{M} - \tilde{\boldsymbol{M}})(\boldsymbol{M} + \tilde{\boldsymbol{M}})\boldsymbol{\Lambda}_0^\top + 2\boldsymbol{\Lambda}_0(\boldsymbol{M} - \tilde{\boldsymbol{M}})\boldsymbol{S}_0^{1/2}\boldsymbol{\Lambda}_0^\top &= 0, && (25)\\
\boldsymbol{\Lambda}_0(\boldsymbol{M} - \tilde{\boldsymbol{M}})\left(\boldsymbol{M} + \tilde{\boldsymbol{M}} + 2\boldsymbol{S}_0^{1/2}\right)\boldsymbol{\Lambda}_0^\top &= 0. && (26)
\end{aligned}$$

Since $\boldsymbol{\Lambda}_0$ is full rank and assuming $\boldsymbol{S}_0 \succ 0$, for the equality to hold, we should have $\boldsymbol{M} = \tilde{\boldsymbol{M}}$ or $\boldsymbol{C} + \tilde{\boldsymbol{C}} + 2\boldsymbol{I}_r = 0$. The second equality is ruled out by assumption that $\boldsymbol{C}$ is positive. Hence, $\boldsymbol{C} = \tilde{\boldsymbol{C}}$ leading to a contradiction. Therefore, there exists no observationally equivalent $\tilde{\boldsymbol{C}}$. ■

References

Almuzara, M., Baker, K., O'Keeffe, H., and Sbordone, A. (2023). The New York Fed staff nowcast 2.0. *New York Fed Staff Nowcast Technical Paper.–2023*.

Anderson, T. W., Rubin, H., et al. (1956). Statistical inference in factor analysis. In *Proceedings of the third Berkeley symposium on mathematical statistics and probability*, volume 5, pages 111–150.

Antolín-Díaz, J., Drechsel, T., and Petrella, I. (2024). Advances in nowcasting economic activity: The role of heterogeneous dynamics and fat tails. *Journal of Econometrics*, 238(2):105634.

Artemova, M. (2025). An order-invariant score-driven dynamic factor model. *Journal of Econometrics*, 251.

Artemova, M., Blasques, F., van Brummelen, J., and Koopman, S. J. (2022). Score-driven models: Methodology and theory. In *Oxford Research Encyclopedia of Economics and Finance*, pages 1–33. Oxford University Press.

Aruoba, S. B., Diebold, F. X., and Scotti, C. (2009). Real-time measurement of business conditions. *Journal of Business & Economic Statistics*, 27(4):417–427.

Bai, J. and Li, K. (2012). Statistical analysis of factor models of high dimension. *The Annals of Statistics*, 40(1):436–465.

Bai, J. and Wang, P. (2015). Identification and Bayesian estimation of dynamic factor models. *Journal of Business & Economic Statistics*, 33(2):221–240.

Blasques, F., Francq, C., and Laurent, S. (2023). Quasi score-driven models. *Journal of Econometrics*, 234(1):251–275.

Blasques, F., Koopman, S. J., Łasak, K., and Lucas, A. (2016). In-sample confidence bands and out-of-sample forecast bands for time-varying parameters in observation-driven models. *International Journal of Forecasting*, 32(3):875–887.

Blasques, F., van Brummelen, J., Gorgi, P., and Koopman, S. J. (2024). A robust Beveridge–Nelson decomposition using a score-driven approach with an application. *Economics Letters*, 236:111588.

Blasques, F., van Brummelen, J., Koopman, S. J., and Lucas, A. (2022). Maximum likelihood estimation for score-driven models. *Journal of Econometrics*, 227(2):325–346.

Buccheri, G., Corsi, F., and Dzuverovic, E. (2024). From rotational to scalar invariance: Enhancing identifiability in score-driven factor models. *arXiv preprint arXiv:2412.01367*.

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

Creal, D., Koopman, S. J., and Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5):777–795.

Crone, T. M. and Clayton-Matthews, A. (2005). Consistent economic indexes for the 50 states. *Review of Economics and Statistics*, 87(4):593–603.

Durbin, J. and Koopman, S. J. (2012). *Time series analysis by state space methods*, volume 38. Oxford University Press.

Harvey, A. C. (1990). Forecasting, structural time series models and the Kalman filter.

Harvey, C. R. (1988). The real term structure and consumption growth. *Journal of financial Economics*, 22(2):305–333.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45.

Koopman, S. J., Lucas, A., and Scharth, M. (2016). Predicting time-varying parameters with parameter-driven and observation-driven models. *Review of Economics and Statistics*, 98(1):97–110.

Koopman, S. J., Lucas, A., and Zamojski, M. (2017). *Dynamic term structure models with score-driven time-varying parameters: estimation and forecasting*. Economic Institute.

Lange, R.-J., van Os, B., and van Dijk, D. (2024). Implicit score-driven filters for time-varying parameter models. *Preprint. https://ssrn. com/abstract*, 4227958.

Marcellino, M. (2006). Leading indicators. *Handbook of economic forecasting*, 1:879–960.

Ng, S. (2021). Modeling macroeconomic variations after COVID-19. Technical Report w29060, National Bureau of Economic Research.

Optima Capital Management (2023). Economic indicators in tug of war: Leading vs coincident. `https://www.optimacapitalmgt.com/insights/economic-indicators-in-tug-of-war`. Accessed: 2025-09-13.

Rudebusch, G. D. and Williams, J. C. (2009). Forecasting recessions: The puzzle of the enduring power of the yield curve. *Journal of Business & Economic Statistics*, 27(4):492–503.

Smetanina, E. (2017). Real-time GARCH. *Journal of Financial Econometrics*, 15(4):561–601.

State Street Global Advisors (2025). Leading economic indicator divergence. `https://www.ssga.com/us/en/institutional/insights/weekly-market-update-16-may-2025`. Accessed: 2025-09-13.

Stock, J. H. and Watson, M. W. (1989). New indexes of coincident and leading economic indicators. *NBER macroeconomics annual*, 4:351–394.

Stock, J. H. and Watson, M. W. (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, 97(460):1167–1179.

Stock, J. H. and Watson, M. W. (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162.

Zou, X., Lin, Y., and Lucas, A. (2025). Improving score-driven density forecasts with an application to implied volatility surface dynamics. Technical report, Tinbergen Institute Discussion Paper.