Essays on the Development and Application of Score-Driven Models in Finance and Macroeconomics

Tibor Pál

Università degli Studi di Salerno, DISES

Supervisor: Prof. Giuseppe Storti

May 2026

Contents

Introduction

Economic and financial time series seldom exhibit linear behavior. Complex social structures and interdependencies inherent in macroeconomic and financial dynamics pose significant challenges in studying the underlying phenomena. With advances in economic analysis, it was recognized early that the behavior of macroeconomic and financial variables is conditional on the magnitude, sign, and timing of disruptions, frequently exhibits heavy tails, and transitions through regimes in which the underlying mechanisms undergo substantial changes. Capturing such nonlinearities has long been a central concern within statistical methodology, fostering sustained innovation in applied econometrics over recent decades.

In response to the growing demands arising from developments in finance and economics, two distinct specifications emerged to address *time-varying parameters* (TVPs) in time-series analysis and econometrics, initially classified by Cox (1981) as *observation-driven* and *parameter-driven* models, respectively.

In parameter-driven models, TVPs follow a latent stochastic process with its own error term. As a result, parameter innovations are not deterministic functions of past observations, and parameters are only indirectly linked to the data. This structure generally renders the likelihood function unavailable in closed form. Estimation is therefore computationally involved and typically relies on simulation-based methods. Prominent examples include stochastic volatility models (e.g., Taylor, 1994; Kim et al., 1998) and stochastic intensity frameworks such as those proposed by Bauwens and Hautsch (2006).

In observation-driven models, by contrast, TVPs evolve deterministically as functions of lagged dependent variables together with contemporaneous and lagged exogenous variables. While this specification yields an unconditionally nonlinear data generating process (DGP), the TVPs remain perfectly predictable one step ahead, conditional on past information. This property allows the likelihood function to be expressed in closed form via the *prediction error decomposition* making estimation and inference more straightforward.

The more tractable estimation process, combined with relatively low computational cost and ease of interpretation, has made observation-driven models a popular choice in financial econometrics and statistics. The recognition of volatility clustering motivated the development of the autoregressive conditional heteroskedasticity (ARCH) model by Engle (1982) and its generalization, the GARCH model, by Bollerslev (1986). These contributions played a central role in the widespread adoption of observation-driven models and stimulated further innovations, including autoregressive conditional duration (Engle and Russell, 1998), the dynamic conditional correlation model (Engle, 2002), Poisson count models (Davis et al., 2003), dynamic copula models (Patton, 2006), and the time-varying quantile model of Engle and Manganelli (2004).

The introduction of the Generalised Autoregressive Score (GAS) framework by Creal et al. (2008) represents a major advancement in the evolution of observation-

driven models. The central innovation of the GAS framework lies in its use of the scaled score (s_t) of the likelihood function as the driving force behind parameter dynamics. This approach provides a unified and coherent framework for incorporating time-varying parameters across a broad class of nonlinear models, encompassing and extending many existing observation-driven specifications. In addition, the GAS framework has opened new research avenues by extending observation-driven modeling to unobserved components, multivariate point processes with time-varying parameters, dynamic copulas, and higher-order conditional moments. Owing to its generality and flexibility, GAS has become a reference methodology for analyzing dynamic structures in financial econometrics.

To enhance economic and statistical interpretation of the *score-driven framework*, it is useful to relate it to the classical volatility models discussed above. Specifically, it can be shown that under Gaussian assumptions, the GAS(1,1) specification nests the standard GARCH(1,1) model[1]:

$$\nabla_t = \frac{\partial \log p(y_t \mid \sigma^2_{t-1})}{\partial \sigma^2_{t-1}} = -\frac{1}{2\sigma^2_{t-1}} + \frac{y_t^2}{2\sigma^4_{t-1}} \qquad S_{t-1} = I^{-1}_{t-1} = 2\sigma^4_{t-1}$$

$$s_t = S_{t-1}\nabla_t = y_t^2 - \sigma^2_{t-1}$$

$$\begin{aligned}\sigma^2_t &= \omega + A_0 s_t + B_1 \sigma^2_{t-1} = \omega + A_0(y_t^2 - \sigma^2_{t-1}) + B_1\sigma^2_{t-1} \\ &= \omega + A_0 y_t^2 + (B_1 - A_0)\sigma^2_{t-1} \\ &\equiv \omega + \alpha y_t^2 + \beta \sigma^2_{t-1} \qquad \alpha = A_0, \quad \beta = B_1 - A_0\end{aligned}$$

At the same time, this equivalence implies that score-driven dynamics are closely tied to the information content of the assumed likelihood, which may become overly restrictive under misspecification. This limitation motivates the methodological advancements developed in this thesis. Specifically, the thesis focuses on the development and refinement of score-driven models to accommodate complex and highly nonlinear economic and financial DGPs, where standard GAS models may prove insufficient. Beyond their common methodological foundation, the chapters are unified by a focus on volatility, uncertainty, and unexpected shocks, with particular emphasis on episodes such as the Global Financial Crisis (GFC) and the COVID-19 pandemic.

The thesis takes the reader from univariate financial volatility modeling, through multivariate state-space analysis with the Kalman filter (Kalman, 1960), to dynamic quantile modeling with less restrictive distributional assumptions. Both univariate and multivariate time-series settings are considered, using parametric and semiparametric approaches.

Chapter 1 focuses on financial applications and introduces the Realized Dynamic Score Exponential GARCH (RDS-EGARCH) model with heavy-tailed and skewed distributions. The proposed framework extends the score-driven literature by combining a joint return–realized measure specification with flexible leverage structures and a two-component volatility dynamics designed to capture persistence and long-memory features.

Empirically, the proposed framework and its specifications improve the joint modeling of returns and volatility, strengthen the representation of asymmetric responses, and provide superior in-sample fit. In out-of-sample forecast evalua-

[1]The notation follows Creal et al. (2008).

tion, they deliver competitive and often improved forecasts of Value-at-Risk and Expected Shortfall across multiple return series.

At the same time, the out-of-sample evaluation does not indicate uniform dominance across all loss functions, assets, and tail probabilities. In particular, differences across models are less pronounced under loss functions that primarily focus on violation frequencies, and the gains from more flexible specifications are partly asset-dependent. Importantly, the Model Confidence Set analysis (Hansen et al., 2011) provides more decisive evidence, showing that the baseline RDS-EGARCH specification consistently outperforms the benchmark beta-skew-t-EGARCH variants.

Chapter 2 shifts the focus to macroeconomics by estimating the natural rate of interest in the United States (US), commonly referred to as r-star. Motivated by the econometric challenges posed by the extreme volatility shocks during the COVID-19 pandemic, the study introduces a score-driven updating mechanism in which volatility persistence becomes fully state dependent. The proposed framework, embedded within a score-driven state-space model, is more effective than the benchmark Laubach–Williams framework (Laubach and Williams, 2003) at disentangling transitory and highly persistent shocks, thereby delivering more accurate and precise estimates of r-star.

Despite these improvements, some limitations remain. Importantly, the divergence from the benchmark estimates is largely driven by the estimated "other factors" component, whose economic interpretation remains ambiguous and whose identification is subject to considerable uncertainty. The relatively low-frequency nature of the data and the delayed availability of key macroeconomic indicators may also limit the timeliness of the estimates. In addition, the Phillips curve slope parameter is estimated with substantial downside uncertainty, limiting the ability to draw reliable inference on the time variation of the Phillips curve. These considerations suggest that further work on refining the identification of latent components and incorporating higher-frequency information could enhance both the robustness and real-time applicability of the framework.

Motivated by the inconclusive evidence on the Phillips curve slope parameter, Chapter 3 investigates the conditional distribution of US inflation within a New Keynesian Phillips curve framework using a dynamic multiple quantile model with quasi score-driven, time-varying parameters. The proposed specification jointly models conditional quantiles and allows the Phillips curve slope to vary both over time and across the distribution, capturing quantile-specific nonlinearities and evolving inflation dynamics.

The empirical results reveal pronounced time variation and substantial heterogeneity in the Phillips curve relationship. The estimated slopes are asymmetric across the distribution, with the slope generally steepest in the lower quantiles and progressively flattening toward the upper tail. Time variation is concentrated primarily in the lower and middle parts of the distribution, implying a time-varying Phillips curve shape.

The findings highlight the importance of expectation formation for identification. Short-run expectations generate richer distributional dynamics, while long-run expectations attenuate higher-order moments but amplify time variation in the cross-quantile Phillips curve profile. Episodes of flattening, particularly in the lower tail, are consistent with the *missing disinflation* following the Great Recession and the post-pandemic period. Importantly, these results, especially the consistently flat slope in the upper tail, are robust across specifications.

Chapter 1

Modeling Volatility in Finance: The Realized Dynamic Score Exponential GARCH

This chapter is based on joint work by Tibor Pál, Giuseppe Storti, and Kok Haur Ng, entitled "Realized Dynamic Score Exponential GARCH", currently under development as a working paper.

Abstract

This chapter introduces the Realized Dynamic Score Exponential GARCH model with heavy-tailed, skewed conditional distributions for the joint modeling of returns and a realized range-based volatility measure. Beyond embedding the score-driven framework within volatility dynamics, a key feature of the model is an enhanced leverage effect, with leverage functions specified in both quadratic and score-driven forms and incorporated into both the measurement and volatility equations. The score-driven dynamic relationship between returns, realized range, and future volatility, together with flexible error distributions, improves empirical fit relative to the beta-skew-t-EGARCH specification. In addition, a two-component structure is introduced to capture persistent volatility dynamics and long-memory behavior. Empirical analysis based on the S&P 500 and two individual assets return series indicates that the proposed model provides superior in-sample fit and competitive, often improved, out-of-sample VaR and ES forecasts. These results reflect an improved ability to capture volatility dynamics, particularly during periods of market turbulence.

1.1 Introduction

Modeling volatility is a central theme in financial econometrics, as understanding volatility dynamics is fundamental to asset pricing, risk management, and forecasting. Volatility clustering, the key stylized fact that motivated the development of the canonical autoregressive conditional heteroskedasticity (ARCH) and generalized autoregressive conditional heteroskedasticity (GARCH) models (Engle, 1982; Bollerslev, 1986), is naturally captured within the baseline formulations of this

modeling framework. However, these models often provide a limited account of other empirical features of financial returns, such as heavy tails, skewness, leverage effects, and long memory in volatility, particularly under non-standard market conditions. Moreover, rapid shifts in volatility can cause standard GARCH-type models to adjust slowly to new volatility regimes.

This study aims to address these challenges jointly. To this end, we introduce a variant of the realized GARCH model that employs score-driven updating dynamics to jointly link the return process and realized range volatility, where volatility dynamics are driven by the scores from both the return and measurement equations. We refer to this specification as the Realized Dynamic Score Exponential GARCH (RDS-EGARCH) model.

The proposed framework makes several contributions. First, it jointly exploits information from returns and realized range-based volatility within a unified score-driven updating mechanism. Second, it incorporates a two-component volatility structure to more effectively capture persistence. Third, it extends the modeling of leverage effects by including both quadratic and score-driven terms in the return and measurement equations. Finally, it combines these features with flexible heavy-tailed and skewed conditional distributions, enhancing robustness to extreme observations.

Early contributions in the literature propose a broad range of model extensions that address specific empirical features of volatility dynamics, but typically in isolation. Heavy tails were incorporated into GARCH models by Bollerslev (1987) and later generalized to allow for asymmetric tail behavior by Hansen (1994). Recognizing the asymmetric impact of returns on conditional volatility, leverage effects were introduced through the EGARCH model of Nelson (1991) and the GJR-GARCH model of Glosten et al. (1993), among others. Persistence in volatility was further addressed by introducing dynamic component structures, most notably in the seminal work of Ding and Granger (1996), which showed how component-based models can approximate long-memory behavior. Relatedly, long-memory dynamics were formalized explicitly within the fractionally integrated GARCH (FIGARCH) framework of Baillie et al. (1996). However, component models, such as that of Ding and Granger (1996), provide a more tractable tool for modeling persistent volatility dynamics, overcoming the statistical difficulties implied by the use of fractionally integrated dynamic models.

Although these frameworks provide essential building blocks for both researchers and practitioners, they adapt slowly to changes in volatility during periods of market stress. This limitation primarily stems from the reliance of the volatility dynamics on a reduced set of information. The development of the score-driven framework (Creal et al., 2008) marked an important milestone in this respect. Specifically, when the dynamics depend on the conditional score rather than specific conditional moments, and the score is derived from a heavy tailed density, the resulting volatility update reacts more strongly and quickly to large but admissible shocks under that density. Moreover, Harvey and Sucarrat (2014) showed that combining score-driven dynamics with a skew-t distribution yields additional robustness by downweighting observations that would be considered outliers under Gaussianity.

Despite the substantial advances offered by the score-driven framework, its performance remains constrained when the updating equation relies solely on an individual signal, as a single return observation conveys limited information about the prevailing volatility. Consequently, as noted by Andersen et al. (2003), return-

based filters often fail to capture abrupt shifts in the true volatility level. With the emergence of realized volatility measures based on high-frequency data, Hansen et al. (2012) introduced the realized GARCH framework for modeling returns and volatility jointly. By incorporating a measurement equation that links the realized measure to the conditional variance, the framework expands the information set used to update volatility dynamics.

Building on evidence that realized measures are often distorted by jumps (Opschoor et al., 2018), Harvey and Palumbo (2023) extended the score-driven methodology to modeling realized volatility. Using members of the exponential generalized beta of the second kind (EGB2) distribution family, the model accommodates these distortions and achieves robust updating through the conditional score. Although this approach provides a flexible and practical framework for modeling realized volatility on its own, it does not utilize information embedded in the return process. In addition, a model for realized volatility alone cannot generate tail risk forecasts in a fully coherent manner unless embedded in a joint return–volatility specification, which limits its applicability in risk-management settings.

Against this background, the Realized Dynamic Score Exponential GARCH (RDS-EGARCH) model is proposed to jointly model asset returns and realized range volatility. The score-driven framework provides a systematic approach to updating volatility dynamics based on the score of the conditional density, allowing for closed-form likelihood and efficient estimation.

Within this framework, leverage effects, persistence, and distributional flexibility are incorporated through augmented score-driven terms, a two-component volatility structure, and heavy-tailed skewed conditional distributions.

Finally, we demonstrate the empirical relevance of the proposed model through applications to the Standard & Poor's (S&P) 500 index as well as individual assets, namely AAPL and BAC. The results indicate that our specification improves in-sample fit and delivers competitive, often enhanced, out-of-sample performance in forecasting value-at-risk (VaR) and expected shortfall (ES) relative to benchmark beta-skew-t-EGARCH variants. Taken together, these findings highlight the benefits of embedding realized measures and flexible error distributions within a score-driven EGARCH framework for financial volatility modeling.

The remainder of the chapter is organized as follows. Section 1.2 presents the statistical framework of the proposed model, focusing on volatility dynamics with leverage effects, score-driven updating, and estimation. Section 1.3 reports the empirical analysis and model comparison, with particular emphasis on forecast evaluation. Section 1.4 concludes the chapter.

1.2 The statistical framework

The proposed model builds on the realized GARCH framework of Hansen et al. (2012). This framework jointly models returns and realized volatility, linking the realized measure to the latent conditional volatility through a measurement equation. As a result, it explicitly accounts for measurement errors in realized measures and allows conditional volatility to react more promptly and smoothly to information extracted from high-frequency data. For ease of presentation, we will start with a simple baseline version of the RDS-EGARCH model and then gradually deepen our analysis of its components.

The structure of the baseline RDS-EGARCH model takes the form

$$r_t = \alpha_t z_t = \exp(\lambda_t) z_t, \tag{1.1}$$
$$y_t = \omega_x \alpha_t^{\rho} \varepsilon_t = \omega_x \exp(\rho \lambda_t) \varepsilon_t, \tag{1.2}$$

where r_t is the log-return of an asset at time t, that is $r_t = \ln(P_t) - \ln(P_{t-1})$, where P_t is the price at time t, y_t is a realized volatility measure entering the measurement equation defined in Eq. (1.2), $\alpha_t = \exp(\lambda_t)$ is a time-varying scale parameter that is assumed to be a function of the information set up to time $t-1$, $\mathcal{I}_{t-1}$.

In the baseline specification of the RDS-EGARCH model, the dynamics are driven by the two mutually independent error terms, $z_t \sim$ i.i.d. $F(\mu_z)$, and $\varepsilon_t \sim$ i.i.d. $G(\mu_\varepsilon)$, where $F(\mu)$ denotes a distribution with mean μ and G is a distribution with positive support. For ease of exposition, we impose $\mathbb{E}[r_t \mid \mathcal{I}_{t-1}] = 0$, consistent with treating z_t as a zero mean innovation.

The parameters ω_x and ρ are scaling parameters, whose role is to align the level of the realized measure with the latent volatility. Namely, they correct for the systematic scaling bias (e.g., Christensen and Podolskij, 2007) of the realized measure and allow to isolate the dynamics in λ_t, ensuring that ε_t reflects only pure innovations. Although we do not explicitly implement bias-corrected estimators as in Christensen et al. (2009), the scaling parameters, together with the score-driven updating mechanism, can partially account for persistent distortions arising from market microstructure noise in the level of the realized measure. This does not substitute for explicit bias correction, and incorporating adjusted realized measures represents a natural extension for future research.

The following subsections introduce the dynamic representation of the volatility and present the additional specifications that lead to the final version of the model.

1.2.1 Heavy tails and skewness

This section provides insight on the distributional framework underline the RDS-EGARCH model presented in Eqs. (1.1)–(1.2). Our settings allow to account for the presence of heavy tails and skewness in the conditional distribution of returns, which is a well known stylized fact. To this aim, the returns innovations z_t are assumed to follow a skewed Student's t distribution, $z_t \sim$ i.i.d. Skew-$t(\nu, \gamma)$ where ν ($\nu > 2$) is the degrees of freedom parameter and γ is an additional parameter controlling skewness. More specifically, the skewed t distribution is obtained through contamination of a $t(\nu)$ distribution following the methodology proposed by Fernández and Steel (1998): given a variable z whose density $f(z)$ is continuous, unimodal and symmetric around zero, skewness is achieved by introducing inverse scale factors in the positive and the negative orthant (see Fernández and Steel, 1998, for more details). The resulting distribution is able to account for both heavy tails and skewness and its density is given by

$$f(z_t \mid \gamma) = \frac{2}{\gamma + \gamma^{-1}} f\left(\frac{z_t}{\gamma^{\mathrm{sgn}(z_t)}}\right), \tag{1.3}$$

where z_t becomes uncentered, with $\gamma \in (0, \infty)$ governing the degree of skewness. Symmetry occurs at $\gamma = 1$, while values of $\gamma < 1$ and $\gamma > 1$ indicate negative and positive skewness, respectively.

As observed by Giot and Laurent (2003), the conditional expectation of the demeaned return

$$\mathbb{E}_{t-1}[r_t] = \mu_z \exp\big(\lambda_{t|t-1}\big), \tag{1.4}$$

where μ_z is the uncentered mean, is not constant. Consequently, r_t cannot be martingale difference (MD). To restore the MD property, following Harvey (2013, Sec. 4.1, p. 146), the model is reformulated as

$$r_t = -\mu_z \exp(\lambda_{t|t-1}) + \exp(\lambda_{t|t-1}) z_t, \tag{1.5}$$

where, given the skew t distribution assumed in the return process, μ_z is defined as

$$\mu_z = \sqrt{\nu}\left(\gamma - \frac{1}{\gamma}\right)\frac{\mathrm{B}\left(\frac{\nu-1}{2}, \frac{1}{2}\right)}{\pi} = \sqrt{\nu}\left(\gamma - \frac{1}{\gamma}\right)\frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi}}. \tag{1.6}$$

In the measurement equation, we use the Burr (or Singh Maddala) distribution to match the positive support of the observations. The Burr distribution is a special case of the generalized beta of the second kind (GB2) density, which allows for fat-tailed behavior through its two shape parameters. Using a fat-tailed distribution is important to accommodate the outliers commonly observed in high-frequency data. Our choice for the Burr distribution is also supported by Harvey and Palumbo (2023) as they identify it as one of the best performing specifications for modeling realized volatility in a score-driven settings. The standard form of GB2, as given in Kleiber and Kotz (2003), is

$$f_{t-1}(y_t; \alpha_t, a, p, q) = \frac{a(y_t/\alpha_t)^{ap-1}}{\lambda_t \,\mathrm{B}(p,q)\,\left[(r_t/\alpha_t)^a + 1\right]^{p+q}}, \qquad 0 < y_t < \infty, \quad \alpha_t, a, p, q > 0,$$

where $\alpha_t = \exp(\lambda_t)$ is the scale parameter, a, p, and q are the shape parameters, and $\mathrm{B}(p,q)$ denotes the beta function; the notation f_{t-1} indicates conditioning on information available at time $t-1$. We use the Burr specification obtained by setting $p = 1$ in the GB2 density that results in

$$f_{t-1}(y_t; \alpha_t, a, q) = \frac{aq(y_t/\alpha_t)^{a-1}}{\alpha_t\left[(r_t/\alpha_t)^a + 1\right]^{1+q}}, \qquad 0 < y_t < \infty, \quad \alpha_t, a, q > 0.$$

In order to enhance the flexibility of the model, we work with the exponential Singh–Maddala (ESM) distribution, obtained by applying a logarithmic transformation to the measurement equation. Consequently, Eq. (1.2) is now expressed as

$$x_t = \log \omega_x + \rho\lambda_t + \log \varepsilon_t, \tag{1.7}$$

where $x_t = \log(y_t)$, with conditional location $\log \omega_x + \rho\lambda_t$. Finally, the density function of the ESM distribution takes the form

$$f_{t-1}(x_t; \lambda_t, a, q, \omega_x, \rho) = \frac{aq\,\exp\{a(x_t - \log\omega_x - \rho\lambda_t)\}}{\left[1 + \exp\{a(x_t - \log\omega_x - \rho\lambda_t)\}\right]^{1+q}}, \qquad a, q, \rho > 0.$$

1.2.2 ML Estimation

This section focuses on the ML estimation of the RDS-EGARCH model outlined in Eqs. (1.19)–(1.18). Relying on standard probability arguments and on the assumed independence of the returns and measurement errors, z_t and ϵ_t, the joint conditional distribution of (r_t, x_t) given $\mathcal{I}_{t-1}$ can be factorized as follows

$$f_{rx,t-1}(r_t, x_t) = f_{r,t-1}(r_t) f_{x|r,t-1}(x_t|r_t). \tag{1.8}$$

After taking logs, Eq. (1.8) leads to the following additive decomposition for the conditional log-likelihood

$$\ln f_{rx,t-1}(r_t, x_t) = \ell_{rx,t-1}(r_t, x_t; \theta) = \ell_{r,t-1}(r_t; \theta) + \ell_{x|r,t-1}(x_t; \theta), \tag{1.9}$$

where θ is the vector of static model parameters to be estimated, $\ell_{r,t-1}(r_t;\theta) = \ln f_{r,t-1}(r_t;\theta)$ and $\ell_{x,t-1}(x_t;\theta) = \ln f_{x,t-1}(x_t;\theta)$. Under distributional assumption in Sec. 1.2.1, the returns log-likelihood component $\ell_{r,t-1}(r_t;\theta)$ is given by

$$\begin{aligned}\ell_{r_t} = \ln(2) - \ln\left(\gamma + \frac{1}{\gamma}\right) + \ln\left(\Gamma\left(\frac{\nu+1}{2}\right)\right) - \ln\left(\Gamma\left(\frac{\nu}{2}\right)\right) \\ - \frac{\ln(\pi\nu)}{2} - \lambda_t - \frac{\nu+1}{2}\ln\left(1 + \frac{(r_t + \mu_z \exp(\lambda_t))^2}{\gamma^{2\,\mathrm{sign}(r_t+\mu_z\exp(\lambda_t))}\exp(2\lambda_t)\nu}\right),\end{aligned} \tag{1.10}$$

while the log-likelihood component related to the measurement equation is

$$\ell_{x|r,t-1}(x_t;\theta) = \ln(a) + a\varepsilon_t + \ln(q) - (1+q)\ln(\exp(a\varepsilon_t) + 1)\,. \tag{1.11}$$

The MLE of θ is then defined as the minimizer of the log-likelihood in Eq. (1.9)

$$\hat{\theta} = \arg\max_{\theta}\ \ell_{rx,t-1}(r_t, x_t; \theta).$$

1.2.3 Scale dynamics

Differently from the standard realized GARCH framework, since z_t is not constrained to have a unit variance, the scale parameter α_t can be interpreted as the conditional variance of returns only up to a proportionality factor. Relying on an exponential link function, the dynamic model is then specified for the log-transformed scale λ_t. Furthermore, compared to the standard realized GARCH specification a distinctive feature of the proposed modeling framework is the incorporation of a score-driven updating mechanism. The key components in the scale dynamics are the conditional scores of the likelihood functions, which feed new information into the scale updating mechanism. In addition, as will be shown, the score-driven framework naturally accommodates a leverage channel, through which the conditional score of the return innovation density drives asymmetric effects on volatility.

More precisely, there are three key departures from the general form of the realized GARCH model of Hansen et al. (2012). First, the dynamic specification of λ_t adopts a two-component structure, similar to Harvey and Sucarrat (2014). This component structure allows to reproduce long memory behavior, in line with Ding and Granger (1996) who have shown that a component GARCH structure can generate patterns which are observationally equivalent to long memory. It is worth noting that, in order to guarantee identifiability of the components and avoid the so called exchangeability problem, it is necessary to introduce constraints on their persistence. These constraints allow to refer to the two components as the long and short memory components, respectively.

The volatility level is therefore expressed as a linear combination of the two components, each governed by its own autoregressive (AR) dynamics. This structure enables the model to capture both highly persistent movements and short-term rapid responses to new information. Modeling the dynamics of long- and short-term volatility separately has some obvious advantages, including a clearer characterization of the leverage effect, as it will be later discussed.

Second, instead of including the realized measure directly in the GARCH equation, the model uses the conditional score of the measurement error density to drive the volatility updates.

Third, the dynamic model for λ_t can be further generalized to include an additional leverage term, which can generate dynamic asymmetric effects on volatility. This can be technically achieved by adding to the specification of the short memory component a term that depends on the sign of past returns as well as on the score of the conditional returns log-likelihood.

The resulting (baseline) specification for λ_t is

$$\lambda_t = \omega_\lambda + \lambda_{1,t} + \lambda_{2,t}, \tag{1.12}$$

$$\lambda_{1,t+1} = \phi_1 \lambda_{1,t} + \kappa_{11} u_{r,t} + \kappa_{12} u_{x,t}, \tag{1.13}$$

$$\lambda_{2,t+1} = \phi_2 \lambda_{2,t} + \kappa_{21} u_{r,t} + \kappa_{22} u_{x,t}, \tag{1.14}$$

where $\lambda_{1,t}$ and $\lambda_{2,t}$ characterize the long-term and short-term components, respectively; $u_{r,t}$ and $u_{x,t}$ are the scores of the log-likelihood components implied by the returns and measurement equations, respectively:

$$u_{r,t} = \frac{\partial \ell_{r,t-1}}{\partial \lambda_t}, \tag{1.15}$$

$$u_{x,t} = \frac{\partial \ell_{x,t-1}}{\partial \lambda_t}, \tag{1.16}$$

where θ is a vector collecting the static parameters of the model. For ease of reference, $u_{r,t}$ and $u_{x,t}$ will be referred to as the returns and measurement scores, respectively.

The long-term component, denoted by $\lambda_{1,t}$, typically has an AR coefficient ϕ_1 close to one, or is even set to unity. In contrast, the short-term component, $\lambda_{2,t}$, usually places relatively greater weight on the score innovations κ_{21} and κ_{22} and is combined with a smaller AR coefficient. The model is not identifiable when $\phi_2 = \phi_1$. Therefore, following Harvey and Sucarrat (2014), identifiability and stationarity are ensured by restricting the AR coefficients to satisfy $0 < |\phi_2| < |\phi_1| < 1$. This constraint also prevents the exchangeability of the long- and short-term components.

The two conditional scores, $u_{1,t}$ and $u_{2,t}$, which enter both components together with their associated weight parameters κ_{11}, κ_{12}, κ_{21}, and κ_{22}, collect information embedded in the likelihood densities associated with the return and measurement equations.

1.2.4 Leverage effect

Although introducing skewness can induce a slight leverage effect, Harvey (2013) finds that this effect is modest in magnitude. The model therefore incorporates two distinct mechanisms through which the leverage effect can operate directly.

First, following Harvey (2013), we introduce a leverage term in the volatility dynamics. Consistent with Engle and Lee (1999) and subsequent findings, this is restricted to operate only through the short term dynamics. Thus, the short-term component defined in Eq. 1.14 becomes

$$\lambda_{2,t+1|t} = \phi_2 \lambda_{2,t|t-1} + \kappa_{21} u_{r,t} + \kappa_{22} u_{x,t} + \kappa^* \operatorname{sgn}(-r_t)\,(u_{r,t} + 1), \tag{1.17}$$

where κ^* determines the contribution of the leverage term.

Note that we restrict the leverage function to $u_{r,t}$ to ensure that volatility asymmetry is driven solely by signed return innovations. Since the realized measure is defined on a positive support, its score $u_{x,t}$ reflects magnitude rather than directional shocks. Including $u_{x,t}$ in the leverage term would therefore introduce sign-independent variation and weaken identification of the weight parameter κ^*.

Second, as in Hansen et al. (2012), an additional leverage function enters the measurement equation in a simple quadratic form of $\tau(z_t) = \tau_1 z_t + \tau_2(z_t^2 - 1)$. Consequently, the measurement equation defined in Eq. (1.7) is modified to

$$x_t = \log \omega_x + \rho\lambda_t + \log \varepsilon_t + \tau(z_t). \tag{1.18}$$

As shown below, allowing the leverage effect to operate through multiple channels makes the model more flexible and further improves its ability to capture the dynamic relationship between returns and volatility.

1.2.5 Score-driven dynamics

Within the proposed modeling framework, the volatility dynamics are governed by an updating mechanism driven by the conditional score of the joint likelihood. Specifically, the returns and measurement conditional scores, which extract information from the likelihood components associated with the returns and measurement equations, respectively, enter both the long- and short-term components of volatility dynamics.

In this section, we provide analytical derivations of the returns and measurement conditional scores under the frameworks introduced in Sections 1.2.1, 1.2.3, and 1.2.4. More specifically, we assume a skew t distribution for the returns innovation z_t, a Burr distribution for the measurement error ε_t; the long- and short- term volatility components are specified as in Eq. (1.13) and Eq. (1.17), respectively, while the measurement equation is defined as in Eq. (1.18). Analytically, the general RDS-EGARCH takes the form

$$\begin{aligned}
r_t &= \exp(\lambda_t) z_t = \sigma_t z_t, && (1.19)\\
\lambda_t &= \omega_\lambda + \lambda_{1,t} + \lambda_{2,t}, && (1.20)\\
\lambda_{1,t+1} &= \phi_1 \lambda_{1,t} + \kappa_{11} u_{r,t} + \kappa_{12} u_{x,t}, && (1.21)\\
\lambda_{2,t+1} &= \phi_2 \lambda_{2,t} + \kappa_{21} u_{r,t} + \kappa_{22} u_{x,t} + \kappa^* \operatorname{sgn}(-r_t)\,(u_{r,t} + 1), && (1.22)\\
x_t &= \log \omega_x + \rho\lambda_t + \log \varepsilon_t + \tau(z_t). && (1.23)
\end{aligned}$$

Formally, the returns score is computed as

$$u_{r,t} = \frac{(\nu + 1)\,[r_t^2 + r_t \mu_z \exp(\lambda_t)]}{\nu \exp(2\lambda_t)\gamma^{2\operatorname{sgn}(c_t)} + c_t^2} - 1, \tag{1.24}$$

where the term c_t is given by

$$c_t = r_t + \mu_z \exp(\lambda_t).$$

The measurement score is derived as

$$u_{x,t} = -a \exp(-2\,\lambda_t)\left(2\,\tau_2\, r_t^2 + \tau_1 \exp(\lambda_t)\, r_t - \rho \exp(2\,\lambda_t)\right) b_t, \tag{1.25}$$

where b_t is given by

$$b_t = \frac{q d_t - 1}{d_t + 1},$$

with the term d_t specified as

$$d_t = \exp\left(-\tau_2 \exp(-2\,\lambda_t)\, r_t^2 - \tau_1 \exp(-\lambda_t)\, r_t - \rho\lambda_t - \omega_x + \tau_2 + x_t\right)^a.$$

The analytical expressions for $u_{r,t}$ and $u_{x,t}$ derived in Eq. (1.24) and Eq. (1.25), respectively, make clear that the dynamics of the two components, $\lambda_{1,t}$ and $\lambda_{2,t}$ non-linearly depend, through both $u_{r,t}$ and $u_{x,t}$, on the past squared returns as well as on past values of the realized measure, through x_t. Also, it is worth noting that the score components depend on the return r_t, thus contributing to the estimation of the leverage effect through a third channel, in addition to those highlighted in Section 1.2.4. Therefore, past values of the score contribute to the dynamics of the scale parameter λ_t in a very complex way, leveraging on both low and high frequency information.

1.3 Empirical analysis

1.3.1 Data description and estimated models

The in-sample analysis is based on the high-frequency data for the S&P 500 index over the period from 3 November 2009 to 22 October 2024, comprising a total of $T = 3767$ trading days. The forecasting exercise additionally considers two individual assets, Apple Inc. (AAPL) and Bank of America Corporation (BAC), for which data are available over the period from 5 January 1998 to 25 October 2024.

The realized volatility measure is constructed from intraday one-minute prices observed during standard trading hours. Specifically, we use a volatility estimator based on the realized range (RR), defined as $y_t = \sqrt{RR_t}$, which has been shown to exhibit lower mean-squared error than realized variance, supported by simulation and empirical evidence (Martens and van Dijk, 2007). The estimator aggregates intraday high–low price ranges and applies a scaling factor originally proposed by Parkinson (1980) to obtain a range-based estimate of daily variance.

Formally,

$$RR_t = \frac{1}{4\ln 2} \sum_{i=1}^{I} \left(\ln H_{t,i} - \ln L_{t,i}\right)^2, \tag{1.26}$$

where

$$H_{t,i} = \max_{1\le s\le n_t} \{\ln(P_{t,s})\}, \qquad L_{t,i} = \min_{1\le s\le n_t} \{\ln(P_{t,s})\}.$$

Here, $s = 1, 2, \ldots, n_t$ indexes intraday observations and $t = 1, 2, \ldots, T$ indexes days. The range of $P_{t,s}$ within each interval t measures the dispersion of intraday prices.

Aggregating intraday price dispersion in this way produces a daily realized range measure that complements return-based information on volatility.

Descriptive statistics for S&P 500 index returns and the associated realized range are reported in Table 1.1. The realized range has a mean of 0.98 and a standard deviation of $\bar{\sigma} = 2.45$, indicating substantial variability. Its unconditional distribution is highly right-skewed ($\bar{\zeta} = 15.45$) and exhibits extremely large excess kurtosis ($\bar{\nu} = 399.80$), reflecting the presence of extreme observations. The Jarque-Bera (JB) test strongly rejects the null hypothesis of normality, confirming the highly non-Gaussian nature of the realized range-based volatility measure.

Returns display a mean close to zero ($\bar{\mu} = 0.03$) and a standard deviation of 0.89. The distribution is moderately negatively skewed ($\bar{\zeta} = -0.64$) and leptokurtic ($\bar{\nu} = 9.69$), indicating asymmetry and heavy tails. The JB test again rejects normality, consistent with the well-documented empirical properties of financial return series.

Table 1.1: Summary statistics of S&P 500 index returns and realized range

Variable	$\bar{\mu}$	$\bar{\sigma}$	$\bar{\zeta}$	$\bar{\nu}$	**JB test**
Realized range	0.981	2.452	15.446	399.800	2.49×10^7 (0.000)
Returns	0.026	0.886	-0.641	9.695	7293 (0.000)

Note: Parentheses report p-values.

Overall, these empirical characteristics motivate the adoption of flexible volatility models that jointly account for heavy tails, asymmetry, and the presence of extreme observations in realized range-based volatility measures.

Model evaluation and comparison are conducted across six specifications. The first two are beta-skew-t-EGARCH models: the original formulation of Harvey and Sucarrat (2014), used here as the benchmark and denoted btE(1), and an extension that incorporates the realized range into the short-term volatility dynamics, denoted btE(2). Specifically, we add a new term to the short-term dynamic volatility component of Eq. (1.14), including the realized measure x_t weighted by parameter δ. We consider these specifications to assess the information content and its impact on the forecast performance of the unmodeled realized measure relative to the score-driven volatility update.

The remaining four models correspond to alternative specifications of the RDS-EGARCH model described in Section 1.2. Specifically, the third model, RDS(1), considers the baseline RDS-EGARCH with a one-component structure and no additional leverage channel. The fourth model, RDS(2), introduces a two-leverage structure while retaining a one-component specification. The fifth model, RDS(3), corresponds to the fully specified RDS-EGARCH with a two-component structure and two leverage channels. Finally, the last specification, RDS(4), applies the same realized measure extension as in the alternative beta skew-t-EGARCH framework. This nested modeling scheme, progressing from the most parsimonious RDS-EGARCH specification to more general variants, is designed to facilitate identification and the evaluation of the marginal contribution of the modeling components introduced in Section 1.2. The considered models and their shorthands are summarized in Table 1.2, which provides a compact overview of the specification hierarchy.

Table 1.2: Overview of model specifications

Shorthand	**Model Specification**
btE(1)	beta-skew-t-EGARCH (Harvey and Sucarrat, 2014)
btE(2)	beta-skew-t-EGARCH with RR in the volatility process
RDS(1)	RDS-EGARCH with one-component, single leverage
RDS(2)	RDS-EGARCH with one-component, two leverage channels
RDS(3)	RDS-EGARCH with two-component, two leverage channels
RDS(4)	RDS(3) with RR in the volatility process

1.3.2 In-sample analysis

Table 1.3 reports parameter estimates, model fit, and diagnostic measures for the beta-skew-t-EGARCH and RDS-EGARCH models, based on the specifications in Table 1.2, estimated on demeaned S&P 500 index returns.

Table 1.3: Parameter estimates, model fit, and diagnostic measures

Parameter	**btE(1)**	**btE(2)**	**RDS(1)**	**RDS(2)**	**RDS(3)**	**RDS(4)**
ω_λ	-0.415	-1.039	-0.026	-0.032	-0.580	-0.037
	[0.164]	[0.005]	[0.004]	[0.004]	[0.076]	[0.096]
ω_x			0.510	0.538	0.537	0.530
			[0.011]	[0.013]	[0.013]	[0.014]
ρ			1.031	1.030	1.049	1.025
			[0.029]	[0.029]	[0.030]	[0.030]
Long term components						
ϕ_1	0.995	0.997			0.986	0.988
	[0.004]	[0.005]			[0.004]	[0.003]
κ_{11}	0.026	0.010			0.026	0.009
	[0.009]	[0.004]			[0.005]	[0.002]
κ_{12}					0.026	0.012
					[0.010]	[0.004]
Short term components						
ϕ_2	0.920	0.782	0.945	0.947	0.707	0.877
	[0.014]	[0.047]	[0.004]	[0.003]	[0.004]	[0.043]
κ_{21}	0.047	0.041	0.052	0.056	0.032	0.035
	[0.013]	[0.009]	[0.004]	[0.004]	[0.006]	[0.004]
κ_{22}			0.122	0.090	0.058	0.035
			[0.010]	[0.007]	[0.010]	[0.008]
δ		0.281				0.166
		[0.075]				[0.036]
Shape						
ν	6.220	6.714	10.768	7.280	7.272	6.938
	[0.604]	[0.70]	[1.535]	[0.707]	[0.687]	[0.657]
γ	0.802	0.799	0.848	0.780	0.789	0.792
	[0.018]	[0.019]	[0.023]	[0.018]	[0.019]	[0.008]
a			8.075	8.900	8.999	9.470
			[0.240]	[0.263]	[0.272]	[0.290]
q			0.639	0.578	0.575	0.531
			[0.035]	[0.032]	[0.032]	[0.029]
Leverage						
κ^*	0.074	0.071		0.055	0.058	0.052
	[0.006]	[0.007]		[0.003]	[0.003]	[0.003]
τ_1			-0.011	-0.036	-0.035	-0.040
			[0.006]	[0.005]	[0.006]	[0.005]
τ_2			0.050	0.031	0.032	0.030
			[0.003]	[0.003]	[0.003]	[0.003]
Model fit						
LogL:r_t	-3971.6	-3937.9	-3939.2	-3902.9	-3907.8	-3901.7
LogL:x_t			-315.5	-130.7	-100.8	-78.1
Total LogL	-3971.6	-3937.9	-4254.6	-4033.6	-4008.6	-3979.8
AIC	7959.1	7893.8	8533.3	8093.2	8049.2	7993.7
BIC	8009.0	7949.9	8608.1	8174.2	8148.9	8099.7
K-S test:r_t	0.022	0.020	0.023	0.022	0.020	0.020
	(0.049)	(0.100)	(0.034)	(0.060)	(0.072)	(0.098)
K-S test:x_t			0.012	0.012	0.013	0.010
			(0.637)	(0.617)	(0.459)	(0.838)

Note: btE denotes the beta–skew-t-EGARCH specification of Harvey (2013). RDS refers to the Realized Dynamic Score EGARCH model. Standard errors of parameter estimates are reported in square brackets; LogL denotes the log-likelihood; and the K–S test reports Kolmogorov–Smirnov statistics, with p-values shown in parentheses.

In all specifications, inertia in the long-term component is strong, with ϕ_1 close to unity. Such large AR coefficients are consistent with the long-memory characteristics typically observed in volatility dynamics. Together with their statistical significance, these estimates suggest that the model appropriately identifies the long-term component under the two-component specifications.

Both leverage effects and negative skewness are pronounced across all models. In specifications with an additional leverage channel, the estimated parameter κ^* associated with the score-driven leverage term is consistent with the standard interpretation that volatility increases following large negative returns. Similarly, the leverage function $\tau(z_t)$ in the RDS-EGARCH models exhibits the expected pattern, as reflected in the sign and statistical significance of its coefficients. Notably, moving from RDS(1) to RDS(2), i.e. introducing a second leverage channel, leads to a substantial improvement in model fit. This provides strong support for incorporating leverage effects in both the measurement equation and the volatility dynamics. Regarding distributional asymmetry, the estimated skewness parameter is around 0.8 under all specifications, indicating a left-skewed return distribution with a higher likelihood of large negative outcomes relative to positive ones.

Turning to the model fit, the distributional assumptions for the measurement innovation appear reasonable, as the Kolmogorov–Smirnov (K-S) test does not reject the null hypothesis at the 0.05 significance level in any case. In contrast, for the return disturbances the null is rejected under the baseline specifications (btE(1) and RDS(1)), with p-values of 0.049 and 0.034, indicating mild deviations from the assumed skew-t distribution. Based on the partial log-likelihood for returns, the btE(2) and the RDS-EGARCH models strongly dominate the baseline skew-t-EGARCH, highlighting the gains from incorporating realized measures into return volatility models. Introducing the additional RR component improves the fit in both model classes, as confirmed by the information criteria. For the extended RDS-EGARCH model, however, the improvement arises mainly in the measurement equation, while the gain in the return equation is marginal. Taken together with the strong significance of the weight parameters, these observations have three potential implications. First, the realized measure is a relevant signal for the return process. Second, the score associated with the measurement equation extracts this information more effectively than directly using the unfiltered realized measure. Third, the marginal improvement in the return equation suggests that the informational contribution of these scores to the return dynamics is already near its upper bound.

To further investigate model performance, we conduct forecast evaluation with a variety of backtests in the following section. We base this exercise on both the predicted out-of-sample VaR and ES, which allows us to assess whether the volatility updating channels provided by different specifications lead to significantly improved predictive accuracy, particularly in the tails.

1.3.3 Forecast evaluation

We use the models described in Table 1.2 to forecast ES and VaR for the daily returns of the S&P 500 index and two additional assets. Specifically, we consider Apple Inc. (AAPL) and Bank of America Corporation (BAC) to assess model performance at the individual asset level.

The out-of-sample evaluation is based on the last 1200 observations in the dataset. For the S&P 500, the initial estimation window spans from November

3, 2009, to January 15, 2020, after which one-step-ahead forecasts are generated. For the two additional assets, the estimation window starts on January 5, 1998. The window is then rolled forward until the final forecast date, October 22, 2024, corresponding to the last available observation.

The out-of-sample interval covers several major episodes of market stress. The evaluation begins prior to the first shocks of COVID-19 in February 2020, which produced exceptional uncertainty and extreme market turbulence. It also contains the onset of the Russia–Ukraine war in early 2022, a sudden geopolitical shock that generated sharp movements in equity and energy markets. Additionally, the period encompasses the global inflation surge from the second half of 2021 through 2022, followed by subsequent monetary tightening, characterized by elevated macro uncertainty. In addition, the evaluation window encompasses the US regional banking stress in March 2023, which triggered a brief but intense spike in risk premia. This episode is particularly relevant for Bank of America Corporation (BAC), as heightened financial sector stress amplifies asymmetric responses in volatility, making it a natural setting to assess specifications with two leverage channels. Finally, the analysis ends after the election-driven risk and renewed geopolitical tensions from mid-2024 to October 2024. The presence of these episodes in the forecasting horizon allows a robust comparison of model performance during volatile market conditions, where tail risk forecasting is most relevant.

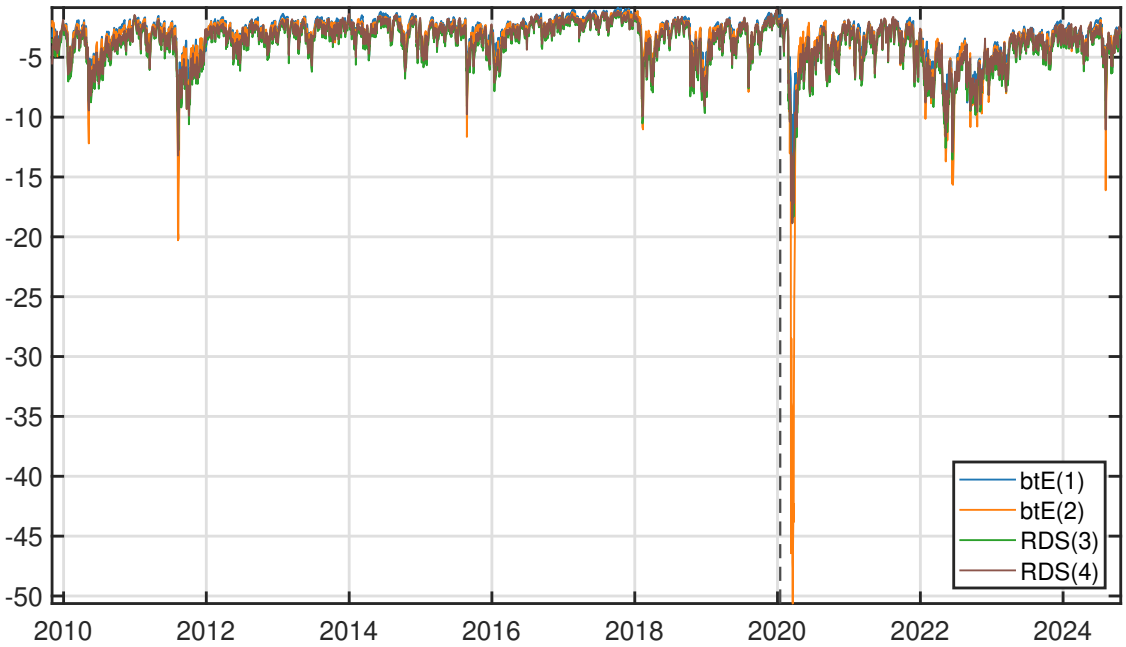

Figure 1.1: Estimated VaR at the 5% level for daily returns of the S&P 500 index over the period November 2009 to October 2024. The dashed vertical line separates the estimation period (left-hand side) from the out-of-sample forecast period (right-hand side). To improve readability, the RDS(1) and RDS(2) models are omitted.

Figure 1.1 displays the fitted VaR at 5% level for the S&P 500 return series for the four considered models. The figure includes the entire sample period, covering both the estimation and evaluation windows. The major volatility episodes align with the pronounced declines in the estimated VaR levels. The figure shows that the average estimated VaR shifted downward in the out-of-sample period, driven primarily by increased uncertainty in the post-COVID era. Although the estimates from the different specifications track each other closely, noticeable discrepancies emerge around unexpected volatility shocks. Notably, the VaR computed by the

beta-skew-t-EGARCH specification with the realized measure is more conservative during these episodes, most clearly during the period affected by the COVID-19 pandemic.

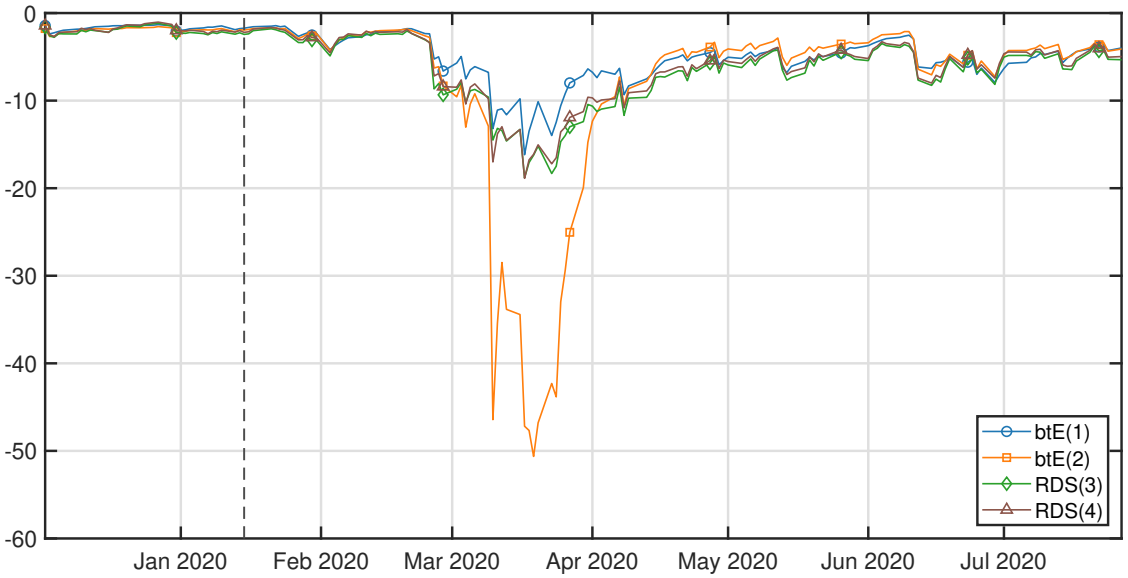

Figure 1.2: Estimated VaR at 5% level for daily returns on the S&P 500 index, over the period December 2019 to July 2020. The dashed vertical line separates the estimation period (left-hand side) from the out-of-sample forecast period (right-hand side). To improve readability, the RDS(1) and RDS(2) models are omitted.

Figure 1.2 zooms in on the interval most affected by the initial COVID-19 shocks to more clearly reveal the variations in the VaR levels. Before the first shock, the estimates are closely aligned, but from the onset of the pandemic, the beta-skew-t-EGARCH specifications diverge markedly from the new models. Interestingly, while the benchmark model yields more optimistic VaR estimates, its alternative specification produces notably more conservative VaR levels than the RDS-EGARCH models. This behavior likely reflects the noisy nature of the realized measure. When the realized measure is not modeled explicitly, its signal disturbance can transmit directly into the volatility forecasts, generating large jumps. In addition, including the realized measure in the volatility dynamics with a static weight parameter limits the flexibility to down-weight such jumps, which the score-driven update would otherwise absorb.

Following the most volatile days, all VaR estimates revert to more moderate levels, yet they remain above their pre-COVID levels and exhibit higher variability. While the deviations are smaller in magnitude, the two beta-skew-t-EGARCH models continue to depart periodically from the other two specifications. However, the direction of these departures is no longer unequivocal, with several episodes showing movements in the opposite direction compared with the most turbulent days. Overall, these observations suggest that directly incorporating the realized measure into the volatility dynamics has a substantial impact on the benchmark model. In contrast, it has only a limited effect on the RDS-EGARCH model.

To further investigate the visually apparent differences in the VaR estimates, we conduct a battery of standard VaR backtests summarized in Table 1.4 on the out-of-sample forecasts. Tables 1.5-1.7 report the results across the considered return series and tail probabilities.

For the S&P 500 (Table 1.5), backtesting at the least conservative level, $\alpha = 0.10$, indicates broadly similar performance across all specifications, with no sys-

Table 1.4: Summary of VaR backtesting procedures

Abbreviation	Full name	Reference
TUFF	Time-until-first-failure test	Kupiec (1995)
POF	Proportion-of-failures test	Kupiec (1995)
CC	Conditional coverage test	Christoffersen (1998)
CCI	Conditional independence test	Christoffersen (1998)
DQ	Dynamic quantile test	Engle and Manganelli (2004)

Table 1.5: VaR backtesting across risk levels for S&P 500 returns

	$\alpha = 0.10$					$\alpha = 0.05$				
	TUFF	POF	CC	CCI	DQ	TUFF	POF	CC	CCI	DQ
btE(1)	0.616	0.567	0.499	0.303	0.902	0.295	0.513	**0.019**	**0.006**	0.141
btE(2)	0.616	0.296	0.346	0.310	0.347	0.295	0.792	**0.033**	**0.009**	*0.081*
RDS(1)	0.616	0.131	0.317	0.902	0.717	0.295	**0.016**	**0.047**	0.558	*0.078*
RDS(2)	0.616	*0.074*	0.106	0.131	0.351	0.295	**0.012**	**0.021**	0.242	**0.046**
RDS(3)	0.616	0.131	*0.083*	*0.101*	0.482	0.295	**0.031**	**0.028**	0.113	*0.081*
RDS(4)	0.616	*0.090*	*0.083*	0.148	0.579	0.295	**0.022**	**0.005**	**0.021**	*0.083*
	$\alpha = 0.025$					$\alpha = 0.01$				
	TUFF	POF	CC	CCI	DQ	TUFF	POF	CC	CCI	DQ
btE(1)	0.136	0.854	0.432	0.200	0.363	**0.048**	0.572	0.722	0.565	0.591
btE(2)	0.136	0.573	0.458	0.265	0.520	*0.100*	0.269	0.438	0.511	0.594
RDS(1)	0.136	0.111	0.273	0.798	0.251	0.278	**0.034**	*0.076*	0.410	**0.034**
RDS(2)	0.136	0.469	0.285	0.159	0.481	*0.100*	0.269	0.438	0.511	0.378
RDS(3)	0.136	0.585	0.339	0.172	0.644	*0.100*	0.269	0.438	0.511	0.755
RDS(4)	0.136	0.714	0.389	0.185	0.740	*0.100*	0.269	0.438	0.511	0.788

Note: TUFF denotes the time-until-first-failure test, POF the proportion-of-failures test, CC the conditional coverage test, CCI the conditional independence test, and DQ the dynamic quantile test. Values for which the null hypothesis is rejected at the 5% significance level are shown in bold, while values rejected at the 10% level are shown in italics.

tematic evidence of misspecification. However, at $\alpha = 0.05$, differences emerge more clearly. Notably, rejections occur in both the conditional coverage (CC) and conditional independence (CCI) tests at the 5% level under the beta-skew-t-EGARCH and the RDS(4) specifications, indicating violations of the joint coverage property and clustering in exceedances. At lower probability levels, $\alpha = 0.025$ and $\alpha = 0.01$, rejection occurrences decline overall, suggesting that all models provide an adequate description of the far tail, with only isolated rejections. Taken together, the evidence points to mild deficiencies in capturing the temporal dependence of violations at intermediate risk levels, rather than systematic failures in unconditional coverage.

For AAPL (Table 1.6), the results are more heterogeneous. At $\alpha = 0.10$, both model classes exhibit frequent rejections in the CC and CCI tests, indicating difficulties in capturing dependence in violations even at relatively moderate risk levels. This pattern weakens at $\alpha = 0.05$, where most specifications pass the unconditional coverage test and only sporadic rejections appear in CC and CCI. Moving further into the tails, however, the RDS-based models show more frequent rejections in the POF and CC tests at $\alpha = 0.025$, suggesting some underestimation of tail risk. At $\alpha = 0.01$, no strong evidence of misspecification is detected across models. Overall, the results indicate that model performance for AAPL is less

Table 1.6: VaR backtesting across risk levels for AAPL returns

	$\alpha = 0.10$					$\alpha = 0.05$				
	TUFF	POF	CC	CCI	DQ	TUFF	POF	CC	CCI	DQ
btE(1)	0.820	0.699	**0.025**	**0.007**	0.202	0.409	0.790	0.419	0.197	0.936
btE(2)	0.820	0.847	**0.020**	**0.005**	**0.078**	0.409	0.419	*0.056*	**0.024**	0.389
RDS(1)	0.820	0.497	0.089	**0.036**	0.260	0.409	0.689	0.428	0.215	0.734
RDS(2)	0.820	0.381	**0.037**	**0.016**	0.115	0.409	0.592	0.429	0.236	0.670
RDS(3)	0.820	0.381	**0.011**	**0.004**	**0.064**	0.409	0.502	0.420	0.257	**0.105**
RDS(4)	0.820	0.561	**0.009**	**0.002**	**0.065**	0.409	0.592	*0.056*	**0.019**	0.283
	$\alpha = 0.025$					$\alpha = 0.01$				
	TUFF	POF	CC	CCI	DQ	TUFF	POF	CC	CCI	DQ
btE(1)	0.192	0.573	0.763	0.637	0.895	*0.068*	0.362	0.617	0.712	0.980
btE(2)	0.192	0.251	0.317	0.322	0.827	*0.068*	0.217	0.442	0.743	0.956
RDS(1)	0.192	**0.049**	0.103	0.410	0.583	*0.068*	0.362	0.617	0.712	0.785
RDS(2)	0.192	**0.009**	**0.026**	0.484	0.301	*0.068*	0.217	0.442	0.743	0.830
RDS(3)	0.192	**0.009**	**0.026**	0.484	0.353	*0.068*	0.217	0.442	0.743	0.783
RDS(4)	0.192	**0.017**	**0.043**	0.459	0.464	*0.068*	0.217	0.442	0.743	0.823

Note: Test statistics follow Table 1.5, with bold and italics denoting rejections at the 5% and 10% significance levels, respectively.

Table 1.7: VaR backtesting across risk levels for BAC

	$\alpha = 0.10$					$\alpha = 0.05$				
	TUFF	POF	CC	CCI	DQ	TUFF	POF	CC	CCI	DQ
btE(1)	0.272	0.628	0.144	*0.057*	0.312	0.123	0.693	0.847	0.674	0.787
btE(2)	0.272	0.923	0.203	*0.075*	0.326	0.123	1.000	0.515	0.249	0.786
RDS(1)	0.272	0.255	0.223	0.192	0.710	0.123	0.362	**0.027**	**0.012**	**0.032**
RDS(2)	0.272	0.255	0.308	0.304	0.787	0.123	0.693	0.144	*0.054*	0.134
RDS(3)	0.272	0.296	0.318	0.273	0.702	0.123	0.693	0.144	*0.054*	0.247
RDS(4)	0.272	0.156	0.200	0.271	0.677	0.123	0.600	0.155	*0.063*	0.164
	$\alpha = 0.025$					$\alpha = 0.01$				
	TUFF	POF	CC	CCI	DQ	TUFF	POF	CC	CCI	DQ
btE(1)	*0.056*	0.854	0.499	0.244	0.318	0.244	0.116	0.278	0.774	0.889
btE(2)	*0.056*	0.853	0.397	0.178	0.299	0.244	0.217	0.442	0.743	0.886
RDS(1)	0.633	0.585	0.491	0.289	0.355	0.244	1.000	0.249	*0.096*	0.139
RDS(2)	0.633	1.000	0.444	0.203	0.485	0.244	0.550	0.160	*0.069*	*0.103*
RDS(3)	0.633	0.714	0.494	0.258	0.357	0.244	0.775	0.280	0.116	0.316
RDS(4)	0.633	0.853	0.397	0.178	0.331	0.244	0.550	0.160	*0.069*	*0.107*

Note: Test statistics follow Table 1.5, with bold and italics denoting rejections at the 5% and 10% significance levels, respectively.

stable across quantiles, with no specification clearly dominating.

For BAC (Table 1.7), the backtesting results are generally more favorable. At $\alpha = 0.10$, all models pass the main coverage tests, with only marginal evidence of dependence in exceedances for the beta-skew-t-EGARCH specifications. At $\alpha = 0.05$, the RDS(1) specification exhibits clear rejections in CC, CCI, and DQ tests, indicating clustering of violations, while the extended RDS specifications with additional leverage channels show improved performance and largely avoid these rejections. This pattern is consistent with the importance of asymmetric volatility responses in financial sector returns, where the inclusion of an additional leverage

channel appears to improve model performance. At lower quantiles, $\alpha = 0.025$ and $\alpha = 0.01$, all models perform satisfactorily, with only isolated and weak rejections.

Overall, the results suggest that differences between model classes are most pronounced at intermediate risk levels. While the RDS-EGARCH framework provides improvements in certain cases, particularly for BAC when additional leverage channels are included, the evidence does not indicate uniform dominance across all assets and quantiles. For the S&P 500, rejection patterns at the 5% level are widespread across all specifications, while for AAPL the results remain mixed with no clear ranking. Instead, model performance appears asset-dependent, with gains from richer leverage structures most evident for financial sector returns and primarily at intermediate tail probabilities.

To complement the diagnostic assessment from the VaR backtests and provide a loss-based comparison of forecasting performance, Table 1.8 reports cross-asset average losses over the out-of-sample period. Two loss functions are considered. First, the standard quantile check loss evaluates the accuracy of VaR forecasts. Second, given the joint availability of VaR and ES forecasts, we report the FZ0 loss function of Patton et al. (2019), which jointly assesses both risk measures.

The quantile loss results consistently favor the RDS-EGARCH specifications, with RDS(1) achieving the lowest loss across all risk levels. However, the differences relative to the alternative RDS variants are small, indicating broadly comparable performance within this class.

When evaluation is based on the joint VaR–ES FZ0 loss, the ranking becomes less clear-cut. Although RDS(1) remains competitive and attains the lowest loss at most quantiles, other specifications, in particular RDS(4), perform similarly and even outperform at $\alpha = 0.05$ level.

Overall, the loss-based comparison indicates that the RDS-EGARCH framework delivers consistently strong performance across assets and risk levels, but the magnitude of improvements over alternative specifications is modest, and model rankings depend on the chosen loss function.

Average losses provide a preliminary view of out-of-sample forecast performance but do not indicate whether the observed differences are statistically significant. Tables 1.9 and 1.10 report Diebold–Mariano (DM) t-statistics for the S&P 500, based on pairwise loss differentials computed using two loss functions. Table 1.9 relies on the quantile check loss, while Table 1.10 is based on the joint VaR–ES FZ0 loss. Loss differentials are defined as the loss of the model in the row minus that in the column, so that negative values indicate superior performance of the row model. Values exceeding 1.64 in absolute value are significant at the 10% level and highlighted in bold.

The results provide only limited evidence of statistically significant differences across models. Under the quantile loss, some significant improvements emerge within the RDS class, with RDS(1) outperforming RDS(2), RDS(3), and RDS(4) at $\alpha = 0.10$, and RDS(4) dominating RDS(3) at multiple risk levels. However, comparisons between RDS and beta-skew-t-EGARCH models rarely reach statistical significance, indicating that the observed loss differences are generally small.

When evaluation is based on the FZ0 loss, the pattern becomes somewhat clearer. In this case, RDS(4) consistently outperforms RDS(2) with statistically significant differences across all risk levels, and also shows significant gains relative to btE(1) at several quantiles. Nonetheless, most other pairwise comparisons remain insignificant, including those involving RDS(1), suggesting that improvements from incorporating ES are not uniformly strong across specifications.

Table 1.8: Out-of-sample average losses (cross-asset mean)

Average quantile loss

	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$
btE(1)	0.257	0.157	0.093	0.047
btE(2)	0.257	0.155	0.092	0.045
RDS(1)	**0.251**	**0.150**	**0.087**	**0.042**
RDS(2)	0.253	0.151	0.088	0.043
RDS(3)	0.253	0.152	0.089	0.043
RDS(4)	0.253	0.151	0.088	0.043

Average FZ0 loss

	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$
btE(1)	0.885	1.100	1.285	1.521
btE(2)	0.867	1.069	1.247	1.455
RDS(1)	**0.848**	1.043	**1.203**	**1.406**
RDS(2)	0.854	1.045	1.211	1.414
RDS(3)	0.855	1.047	1.217	1.422
RDS(4)	0.851	**1.038**	1.206	1.407

Note: The table reports cross-asset average losses (mean across S&P 500, AAPL, and BAC). The upper panel shows the quantile loss, while the lower panel reports the FZ0 loss. The lowest average loss in each column is in bold.

Finally, comparisons between the two beta-skew-t-EGARCH variants reveal no consistent or statistically significant ranking, with DM statistics remaining well below conventional thresholds across all risk levels.

Overall, the DM tests confirm that while RDS-type models tend to yield lower average losses, the statistical evidence in favor of one specific specification is limited and depends on both the loss function and the tail probability considered.

Moving from pairwise comparisons to an overall evaluation of forecast accuracy, we apply the model confidence set (MCS) procedure of Hansen et al. (2011). We consider confidence levels of 90% and 75%, corresponding to exclusion thresholds of 0.10 and 0.25. Table 1.11 reports the resulting average losses and the inclusion frequencies in the superior model sets $\hat{\mathcal{M}}_{90\%}$ and $\hat{\mathcal{M}}_{75\%}$ across assets and risk levels.

For the quantile loss, the results indicate a clear dominance of the RDS specifications. The RDS(1) model is included in the superior set in all cases, achieving the maximum inclusion frequency. RDS(4) also performs strongly, with near-complete inclusion. In contrast, the beta-skew-t-EGARCH models are less robust, particularly btE(1), which is included only sporadically and rarely survives the more stringent 75% confidence level. btE(2) performs better but is still dominated by the RDS variants, especially in the tails. These patterns are largely consistent across all three assets and tend to become more pronounced at more extreme quantiles.

The evidence is even stronger when considering the joint VaR–ES losses based on the FZ0 function. Here, RDS(1) and RDS(4) again achieve full inclusion across all configurations at both confidence levels, while btE(1) is almost entirely excluded from the superior set. btE(2) remains competitive in some cases but does not match the consistency of the RDS models. The inclusion frequencies confirm that the advantage of the RDS framework becomes more pronounced once ES information is incorporated.

The beta-skew-t-EGARCH model augmented with the realized range often per-

Table 1.9: Diebold–Mariano t-statistics on average out-of-sample quantile check loss differences for the S&P 500 index

	$\alpha = 0.10$					
	btE(1)	btE(2)	RDS(1)	RDS(2)	RDS(3)	RDS(4)
btE(1)		0.088	1.672	1.142	0.799	1.245
btE(2)	-0.088		0.619	0.254	0.020	0.114
RDS(1)	**-1.672**	-0.619		**-2.277**	**-1.816**	**-1.676**
RDS(2)	-1.142	-0.254	2.277		-1.167	-0.779
RDS(3)	-0.799	-0.020	1.816	1.167		2.131
RDS(4)	-1.245	-0.114	1.676	0.779	**-2.131**	

	$\alpha = 0.05$					
	btE(1)	btE(2)	RDS(1)	RDS(2)	RDS(3)	RDS(4)
btE(1)		0.504	1.514	1.413	1.340	1.662
btE(2)	-0.504		0.113	-0.029	-0.295	-0.178
RDS(1)	-1.514	-0.113		-0.883	-1.234	-1.006
RDS(2)	-1.413	0.029	0.883		-1.116	-0.772
RDS(3)	-1.340	0.295	1.234	1.116		1.930
RDS(4)	**-1.662**	0.178	1.006	0.772	**-1.930**	

	$\alpha = 0.025$					
	btE(1)	btE(2)	RDS(1)	RDS(2)	RDS(3)	RDS(4)
btE(1)		0.646	1.384	1.363	1.277	1.558
btE(2)	-0.646		-0.084	-0.045	-0.402	-0.249
RDS(1)	-1.384	0.084		0.269	-1.071	-0.700
RDS(2)	-1.363	0.045	-0.269		-1.117	-0.797
RDS(3)	-1.277	0.402	1.071	1.117		1.768
RDS(4)	-1.558	0.249	0.700	0.797	**-1.768**	

	$\alpha = 0.01$					
	btE(1)	btE(2)	RDS(1)	RDS(2)	RDS(3)	RDS(4)
btE(1)		0.758	0.994	1.237	1.544	1.704
btE(2)	-0.758		-0.437	0.006	-0.302	-0.087
RDS(1)	-0.994	0.437		1.137	0.483	1.390
RDS(2)	-1.237	-0.006	-1.137		-0.699	-0.235
RDS(3)	-1.544	0.302	-0.483	0.699		1.799
RDS(4)	**-1.704**	0.087	-1.390	0.235	**-1.799**	

Note: The table reports t-statistics from Diebold–Mariano tests applied to quantile losses over the out-of-sample period January 2020 to October 2024 for the six forecasting models. A positive statistic indicates that the model in the row produces a higher average loss than the model in the column. Absolute values above 1.64, highlighted in bold, indicate significance at the 90% confidence level. Diagonal elements are zero by definition and are omitted for clarity.

forms comparably in the MCS results based on quantile and FZ0 loss functions. This similarity appears somewhat at odds with the pronounced differences in VaR dynamics during volatile periods observed in Figures 1.1-1.2. This is not necessarily contradictory. Standard quantile-based losses are driven by the frequency and magnitude of violations, but are relatively insensitive to the degree of conservatism when no violation occurs. As a result, models that systematically produce more conservative VaR forecasts may avoid penalties from violations while not being

Table 1.10: Diebold–Mariano t-statistics on out-of-sample FZ0 loss differences for the S&P 500 index

	$\alpha = 0.10$					
	btE(1)	btE(2)	RDS(1)	RDS(2)	RDS(3)	RDS(4)
btE(1)		1.507	1.238	1.090	1.385	1.538
btE(2)	-1.507		0.293	-0.038	0.261	0.553
RDS(1)	-1.238	-0.293		-0.883	-0.193	0.476
RDS(2)	-1.090	0.038	0.883		1.261	2.342
RDS(3)	-1.385	-0.261	0.193	-1.261		1.558
RDS(4)	-1.538	-0.553	-0.476	**-2.342**	-1.558	

	$\alpha = 0.05$					
	btE(1)	btE(2)	RDS(1)	RDS(2)	RDS(3)	RDS(4)
btE(1)		1.517	1.043	1.335	1.532	1.674
btE(2)	-1.517		0.138	0.351	0.520	0.825
RDS(1)	-1.043	-0.138		0.329	0.896	1.254
RDS(2)	-1.335	-0.351	-0.329		0.903	2.036
RDS(3)	-1.532	-0.520	-0.896	-0.903		1.550
RDS(4)	**-1.674**	-0.825	-1.254	**-2.036**	-1.550	

	$\alpha = 0.025$					
	btE(1)	btE(2)	RDS(1)	RDS(2)	RDS(3)	RDS(4)
btE(1)		1.460	0.874	1.408	1.541	1.679
btE(2)	-1.460		0.074	0.586	0.684	0.997
RDS(1)	-0.874	-0.074		0.787	1.226	1.395
RDS(2)	-1.408	-0.586	-0.787		0.705	1.839
RDS(3)	-1.541	-0.684	-1.226	-0.705		1.490
RDS(4)	**-1.679**	-0.997	-1.395	**-1.839**	-1.490	

	$\alpha = 0.01$					
	btE(1)	btE(2)	RDS(1)	RDS(2)	RDS(3)	RDS(4)
btE(1)		1.587	0.601	1.541	1.631	1.778
btE(2)	-1.587		-0.133	0.730	0.805	1.148
RDS(1)	-0.601	0.133		1.085	1.392	1.450
RDS(2)	-1.541	-0.730	-1.085		0.716	1.865
RDS(3)	-1.631	-0.805	-1.392	-0.716		1.472
RDS(4)	**-1.778**	-1.148	-1.450	**-1.865**	-1.472	

Note: The table reports t-statistics from Diebold–Mariano tests applied to losses computed with the FZ0 loss function. Further details are given under Table 1.9.

sufficiently penalized for excessive capital allocation.

$$f_1(\varepsilon_{t+1}, VaR_{m,t+1}) = \left| 1 - \left| \frac{\varepsilon_{t+1}}{VaR_{m,t+1}} \right| \right| \tag{1.27}$$

$$f_2(\varepsilon_{t+1}, VaR_{m,t+1}) = \frac{\left(|\varepsilon_{t+1}| - |VaR_{m,t+1}|\right)^2}{|VaR_{m,t+1}|} \tag{1.28}$$

$$f_3(\varepsilon_{t+1}, VaR_{m,t+1}) = |\varepsilon_{t+1} - VaR_{m,t+1}| \tag{1.29}$$

This aspect is consistent with the argument in Caporin (2008), who emphasizes that loss functions focused on violations may fail to capture the opportunity costs associated with overly conservative risk measures. To better understand the

Table 1.11: Model confidence set: S&P 500, AAPL and BAC

Loss: Quantile

Asset	S&P 500				AAPL				BAC				Inclusion Freq.	
α	0.1	0.05	0.025	0.01	0.1	0.05	0.025	0.01	0.1	0.05	0.025	0.01	$\hat{\mathcal{M}}^*_{90\%}$	$\hat{\mathcal{M}}^*_{75\%}$
btE(1)	*0.209*	*0.130*	*0.079*	0.041	***0.278***	*0.167*	0.097	0.047	***0.283***	*0.173*	***0.104***	***0.052***	9	4
btE(2)	*0.216*	***0.133***	*0.082*	0.042	***0.275***	***0.163***	***0.093***	***0.044***	***0.279***	***0.170***	***0.100***	***0.048***	11	9
RDS(1)	***0.203***	***0.125***	***0.074***	***0.036***	***0.273***	***0.160***	***0.091***	***0.044***	***0.278***	***0.166***	***0.096***	***0.047***	12	12
RDS(2)	*0.206*	***0.126***	***0.074***	*0.037*	***0.274***	***0.161***	0.093	0.044	***0.279***	***0.167***	***0.096***	***0.047***	10	8
RDS(3)	*0.207*	***0.126***	***0.076***	0.038	***0.275***	***0.161***	***0.092***	0.044	***0.278***	***0.168***	***0.098***	***0.047***	10	9
RDS(4)	*0.206*	***0.126***	***0.075***	*0.037*	***0.274***	***0.161***	***0.092***	***0.044***	***0.278***	***0.165***	***0.096***	***0.047***	12	10

Loss: FZ0, Patton et al. (2019)

Asset	S&P 500				AAPL				BAC				Inclusion Freq.	
α	0.1	0.05	0.025	0.01	0.1	0.05	0.025	0.01	0.1	0.05	0.025	0.01	$\hat{\mathcal{M}}^*_{90\%}$	$\hat{\mathcal{M}}^*_{75\%}$
btE(1)	*0.683*	0.927	1.146	*1.415*	0.953	1.147	1.303	1.521	*1.020*	1.227	*1.405*	***1.626***	5	1
btE(2)	***0.665***	***0.890***	*1.109*	***1.363***	0.934	1.117	***1.264***	***1.438***	***1.003***	*1.201*	***1.367***	***1.565***	10	8
RDS(1)	***0.631***	***0.855***	***1.036***	***1.261***	***0.923***	***1.098***	***1.245***	***1.428***	***0.989***	*1.175*	***1.329***	***1.529***	12	11
RDS(2)	***0.646***	***0.858***	***1.043***	***1.276***	0.928	1.107	1.263	1.446	***0.989***	*1.169*	***1.327***	***1.521***	8	7
RDS(3)	***0.649***	***0.859***	***1.052***	***1.283***	0.930	1.107	1.262	1.453	***0.986***	*1.175*	***1.338***	***1.529***	8	7
RDS(4)	***0.641***	***0.849***	***1.042***	***1.263***	***0.926***	***1.103***	***1.255***	***1.438***	***0.985***	***1.163***	***1.321***	***1.520***	12	12

Loss: Caporin (2008)

Asset	S&P 500				AAPL				BAC				Inclusion Freq.	
α	0.1	0.05	0.025	0.01	0.1	0.05	0.025	0.01	0.1	0.05	0.025	0.01	$\hat{\mathcal{M}}^*_{90\%}$	$\hat{\mathcal{M}}^*_{75\%}$
							f_1							
btE(1)	*0.557*	0.607	0.672	0.741	0.555	*0.597*	***0.650***	*0.711*	0.570	0.608	0.661	0.722	4	1
btE(2)	*0.554*	0.606	0.673	0.741	***0.547***	***0.592***	***0.647***	***0.707***	***0.561***	0.600	0.652	0.713	6	5
RDS(1)	*0.551*	***0.584***	***0.638***	***0.699***	***0.546***	***0.592***	***0.645***	***0.705***	***0.561***	***0.591***	***0.640***	***0.698***	12	11
RDS(2)	***0.550***	*0.586*	0.649	0.718	0.547	0.596	0.652	0.713	***0.560***	***0.592***	0.644	0.706	4	3
RDS(3)	*0.550*	0.591	0.655	0.723	0.547	0.595	0.652	0.713	***0.559***	***0.591***	0.641	0.700	3	2
RDS(4)	***0.548***	0.591	0.656	0.725	***0.544***	***0.591***	***0.648***	0.710	***0.559***	***0.592***	0.645	0.708	6	6
							f_2							
btE(1)	0.539	0.796	1.180	1.808	0.781	1.104	*1.530*	2.173	0.836	1.179	1.649	2.375	1	0
btE(2)	0.649	0.997	1.458	2.177	*0.749*	***1.070***	***1.487***	***2.107***	0.782	1.114	1.560	2.241	4	3
RDS(1)	0.496	***0.696***	***0.973***	***1.390***	*0.750*	***1.075***	***1.488***	***2.095***	0.758	***1.057***	***1.462***	***2.071***	10	9
RDS(2)	***0.487***	***0.702***	1.027	1.544	0.757	1.106	1.551	2.210	0.754	1.067	1.501	2.169	2	2
RDS(3)	0.498	0.721	1.054	1.584	0.754	1.100	1.543	2.199	0.758	***1.053***	***1.458***	***2.070***	3	3
RDS(4)	0.495	0.722	1.063	1.609	***0.738***	***1.069***	***1.503***	2.156	***0.750***	1.067	1.508	2.191	4	4
							f_3							
btE(1)	1.454	1.891	2.382	3.107	***2.140***	***2.683***	***3.247***	*4.022*	2.110	2.693	3.313	4.179	4	3
btE(2)	1.617	2.123	2.686	3.497	***2.128***	***2.658***	***3.207***	***3.954***	2.080	2.639	3.228	4.040	4	4
RDS(1)	1.414	***1.770***	***2.136***	***2.638***	***2.140***	***2.665***	***3.210***	***3.944***	*2.041*	***2.561***	***3.105***	***3.847***	11	10
RDS(2)	***1.399***	1.783	2.208	2.819	2.155	2.710	3.290	4.077	***2.036***	2.578	3.156	3.961	2	2
RDS(3)	1.413	1.805	2.239	2.862	2.152	2.703	3.280	4.065	***2.027***	***2.553***	***3.101***	***3.843***	4	4
RDS(4)	1.412	1.808	2.252	2.892	***2.126***	***2.665***	***3.234***	4.017	***2.034***	2.578	3.167	3.989	4	4

Note: The table reports average losses from the MCS procedure. Models included in $\hat{\mathcal{M}}^*_{90\%}$ and $\hat{\mathcal{M}}^*_{75\%}$ are shown in italics and bold, respectively.

source of the similar forecasting performance, we employ alternative loss specifications (Eqs. (1.27)–(1.29)) proposed by Caporin (2008), which account not only for violations but also for the distance between forecasts and realizations, thereby mitigating the tendency of standard loss functions to favor overly conservative VaR

forecasts and potentially misidentify the correctly specified model.

Across all variants of the alternative loss functions (f_1–f_3), RDS(1) systematically attains the highest inclusion frequencies, often accompanied by RDS(4). The beta-skew-t-EGARCH models appear in the superior set only intermittently and mainly under less stringent confidence levels or for specific series. This pattern suggests that, once opportunity costs are explicitly incorporated into the evaluation, the relative performance of the RDS-EGARCH models is further strengthened, pointing to a more favorable trade-off between risk coverage and capital allocation.

Overall, the MCS results are broadly consistent with the Diebold–Mariano comparisons and the average loss rankings, but provide more decisive evidence in favor of the RDS-EGARCH specifications. The baseline RDS(1) specification consistently exhibits stronger and more stable forecasting performance across return series, loss functions, and tail probabilities.

1.4 Conclusions

Traditional GARCH models have several well-known limitations, particularly in handling heavy-tailed and skewed distributions, as well as leverage effects and long-memory properties in financial data. These shortcomings can affect the accuracy and reliability of volatility forecasts in financial markets.

To overcome these challenges, this chapter introduced the Realized Dynamic Score EGARCH model. The proposed model employs a score-driven framework for jointly modeling asset returns and a range-based realized volatility measure. Additionally, the model incorporates skewed and heavy-tailed distributions, which are inherent features of financial time series but are often not well captured by traditional GARCH models. The framework also includes a two-component volatility structure and extended leverage channels, which increase flexibility and improve model fit. By exploiting the information contained in the conditional densities through their scores, together with additional enhancements, the model enables more responsive and adaptive modeling of financial volatility.

Empirical analysis based on in-sample fit and out-of-sample VaR and ES forecasts for the S&P 500 and the two considered individual assets indicates that the proposed models outperform beta-skew-t-EGARCH variants. These findings reflect improved model fit and a stronger ability to capture volatility dynamics. While not uniformly dominant, the MCS results provide consistent support for the proposed specification, particularly for the more parsimonious variants, which exhibit stable performance across loss functions and tail probabilities.

Given the longer data availability for the two individual assets, future research could investigate model performance over extended forecasting horizons and across broader market environments. In addition, extending the set of considered assets, particularly to those where asymmetric volatility responses and model flexibility are especially relevant, would provide a more comprehensive assessment of the framework.

Finally, the two-component structure of the model, which captures long-run dynamics, naturally motivates the evaluation of multi-step-ahead forecasts. Assessing such horizons would offer additional insight into the ability of the model to capture persistent volatility components and long-memory features.

Chapter 2

Estimating the R-Star in the US: A Score-Driven State-Space Model with Time-Varying Volatility Persistence

This chapter is based on joint work by Tibor Pál and Giuseppe Storti, entitled "Estimating the R-Star in the US: A Score-Driven State-Space Model with Time-Varying Volatility Persistence" (working paper, 2025).

Abstract

This chapter investigates the dynamics of the natural rate of interest (r-star) in the US using a score-driven state-space model within the Laubach–Williams structural framework. Compared to standard score-driven specifications, the proposed model improves flexibility in variance adjustment by assigning time-varying weights to both the conditional likelihood score and the inertia coefficient in the volatility updating equations. The improved state dependence of volatility dynamics effectively accounts for sudden shifts in volatility persistence induced by highly volatile unexpected events. In addition, allowing time variation in the IS and Phillips curve relationships enables the analysis of structural changes in the US economy that are relevant to monetary policy. The results indicate that the advanced models improve the precision of r-star estimates by responding more effectively to changes in macroeconomic conditions.

2.1 Introduction

Due to the unusual magnitude of economic disturbances generated by the Global Financial Crisis (GFC) and the COVID-19 pandemic, potential nonlinearities in macrovariables have come back into the spotlight. Considering the macroeconomic relevance of the natural rate of interest, or r-star, and the recent development of the statistical toolkit of time-varying parameter (TVP) models, it is surprising that the empirical methodologies employed for its estimation have remained essentially unchanged since the seminal work of Laubach and Williams (2003). We fill this

gap by making the system matrices in the Laubach-Williams methodology (LW, thereafter) time-varying, based on the score-driven state-space framework developed by Delle Monache et al. (2021). In addition, to address model instability due to extraordinary transitory shocks, we propose an extension of the accelerating generalised autoregressive score model (aGAS) of Blasques et al. (2019), in which volatility dynamics become fully state-dependent in a multivariate setting.

The natural rate of interest is a crucial macroeconomic benchmark for investors and central banks, as it provides an estimate of the real short-term interest rate that maintains a constant inflation rate. Therefore, its vital role lies in its information content on the position of the actual real interest rate compared to its natural counterpart. More specifically, when the actual real rate is above the r-star, the economy operates under its potential level, implying deceleration of the inflation rate. Nevertheless, the natural rate of interest is an unobserved variable, a feature that has motivated various methodological approaches and led to considerable estimation uncertainty since the concept was first proposed by Wicksell (1936).

Recognizing the low frequency and time-varying nature of macroeconomic unobservables, the Kalman filter (Kalman, 1960) has been shown to be the optimal choice for data-driven estimation of latent variables, as demonstrated in Staiger et al. (1996), Gordon (1998), and Laubach (2001). By identifying the natural rate of interest through economic theory in a semi-structural framework, the prevailing approach to its real-time estimation has become the LW methodology, where the Kalman filter extracts the permanent or highly persistent changes in the real short-term interest rate consistent with a stable inflation rate. Although the LW framework provides a coherent way to study the unobservable components of the model, the initial study found the r-star estimates to be highly imprecise and prone to substantial real-time measurement error (Laubach and Williams, 2003). Despite the uncertainty around model estimations, the original approach has been applied in numerous studies for different economies and with various extensions (e.g. Garnier and Wilhelmsen, 2009; Holston et al., 2017; Krustev, 2019). However, the global financial crisis (GFC) drew attention to the potential instability in the underlying New Keynesian (NK) structural relationships, casting doubt on the static slope of the model framing IS and Phillips curves. For example, the study by Ball (2011) suggests altering the inflation dynamics in the post-GFC interval. In addition, Rossi et al. (2024) find a substantial change in the slope of the Phillips curve since 1980. While the related quantitative results are not inclusive and remain without unanimous consensus, most studies reject the time-invariant nature of the NK structural relationships.

Concerns about macro-nonlinearities culminated with the onset of the COVID-19 pandemic. Extremely volatile episodes generated by exogenous events made the LW static volatility model less capable of capturing low-frequency changes in the unobservables. In response, Holston et al. (2023) introduced a macroeconometric adjustment to the existing model, specifically applied to the interval affected by the pandemic. While the amended model effectively handles the pandemic-related extreme volatility, the implemented restrictions and the ad-hoc nature of the adjustment limit its universal and real-time applicability.

We address the above issues and estimate the r-star in the US by making the disturbance volatilities and the structural relationships in the LW framework time-varying, governed by the conditional score of the likelihood function. The proposed model is an extension of the class of score-driven state-space models introduced by Delle Monache et al. (2021), incorporating a modified version of the accelerating

GAS (aGAS) of Blasques et al. (2019). We coin the new type of aGAS as augmented accelerating GAS (aaGAS), where, in addition to the time-varying weight of the conditional likelihood score, the autoregressive coefficient can become time-varying as well, making the volatility dynamics fully state-dependent. To perform a thorough assessment of the estimation uncertainty, we combine the inferential setup of Blasques et al. (2016), where only the parameter uncertainty is considered, with a simulation procedure based on the methodology proposed by Hamilton (1986), which allows us to incorporate the impact of filter uncertainty, an aspect particularly relevant to models of unobservable components. The results show that the proposed methodological improvements provide a more effective way to disentangle transitory and permanent shocks, resulting in a more accurate estimation of the state variables. Finally, relying on a purely statistical, data-driven approach, the model is expected to handle and identify unexpected future volatility shocks without the need for major ad hoc econometric adjustments, thus providing a valid alternative empirical approach for the real-time estimation of the natural rate of interest.

The chapter is structured as follows. Section 2.2 outlines the econometric framework, beginning with the time-varying parameter model in Section 2.2.1, followed by the baseline score-driven state-space model in Section 2.2.2. Section 2.2.3 extends the framework by introducing an augmented accelerating score-driven approach for time-varying volatility persistence. Section 2.3 presents the empirical analysis, starting with the data description in Section 2.3.1 and proceeding to the model specification and estimation details in Section 2.3.2. The parameter estimates are discussed in Section 2.3.3, while Sections 2.3.4 and 2.3.5 report the estimated time-varying volatilities and the dynamic IS and Phillips curves, respectively. Section 2.3.6 then examines the estimated state variables, focusing on the r-star and output gap. Lastly, Section 2.4 discusses alternative identification and extensions in the Laubach–Williams framework. Section 2.5 concludes.

2.2 The econometric framework

2.2.1 The structure of the time-varying parameter model

In line with the standard Laubach–Williams (LW) approach, the natural rate of interest is defined as the real interest rate consistent with stable inflation and output at potential, where potential output is the latent trend component of real output and can be interpreted as the level of output consistent with stable inflation. Its econometric identification relies on two structural relationships implied by this definition. Specifically, the intertemporal IS equation links the natural rate of interest to the output gap, defined as the percentage difference between actual and potential output, which is in turn related to inflation through the Phillips curve. Accordingly, the following measurement equations govern the model dynamics and provide the basis for identifying the natural rate of interest:

$$\tilde{y}_t = y_t - y_t^*, \tag{2.1}$$

$$\tilde{y}_t = a_1 \tilde{y}_{t-1} + a_2 \tilde{y}_{t-2} + \frac{a_{\tilde{y},t}}{2} \sum_{j=1}^{2} \left(r_{t-j} - r_{t-j}^* \right) + \varepsilon_{\tilde{y},t} \qquad \varepsilon_{\tilde{y},t} \sim \mathcal{N}(0, \sigma_{\tilde{y},t}), \tag{2.2}$$

$$\pi_t = b_1 \pi_{t-1} + (1 - b_1) \pi_{t-2,4} + b_{\pi,t} \tilde{y}_{t-1} + \varepsilon_{\pi,t} \qquad \varepsilon_{\pi,t} \sim \mathcal{N}(0, \sigma_{\pi,t}), \tag{2.3}$$

where y_t and y_t^* denote the natural logarithm of real output and potential output, respectively, scaled by 100. Similar to Holston et al. (2017) (HLW, thereafter), r_t and r_t^* denote the real federal funds rate and its natural rate (r^*), respectively. π_t is the annualized quarterly consumer price inflation rate, and $\pi_{t-2,4}$ is the average of the second to fourth lags of the inflation rate. The system of Eq. (2.2)-(2.3) represents the NK intertemporal IS and Phillips curve relationships based on the HLW specification with a departure for their respective slope parameters, $a_{\tilde{y},t}$ and $b_{\pi,t}$, which become time-varying in the present model. In the IS equation, the series of output gaps $\tilde{y}_t = y_t - y_t^*$, representing the percentage difference between actual and potential output, is determined by its lagged values, a moving average of the lagged deviation of the real funds rate from its natural rate, and a Gaussian white noise error. In the Phillips curve equation, inflation dynamics π_t are governed by past inflation rates, the lagged output gap $\tilde{y}_t$, and a GWN error process.

For ease of reference, Table 2.1 summarizes the considered variables and their definitions.

Table 2.1: Variable Definitions

Variable	**Definition**
r_t	Real interest rate: federal funds rate minus four-quarter moving average inflation.
r_t^*	Natural rate of interest: real interest rate consistent with stable inflation and output at potential.
y_t	Natural logarithm of real gross domestic product (GDP) scaled by 100.
y_t^*	Natural logarithm of potential output scaled by 100 (level of output consistent with stable inflation).
$\tilde{y}_t$	Output gap, percentage (log) difference between actual and potential output.
π_t	Annualized quarterly growth rate of core personal consumption expenditure (PCE) inflation.

The presence of the two stochastic terms, $\varepsilon_{\tilde{y},t}$ and $\varepsilon_{\pi,t}$, plays a key role in disentangling transitory shocks from persistent structural shocks that affect the dynamics of the state variables. Moreover, it is well known that changing economic conditions are likely to lead to volatility clustering in the dynamics of macro-variables, such as inflation and output. If not correctly modeled, this feature can contaminate the estimates of the state variables and, hence, the structural components of the observed variables. Therefore, we depart from the typical static LW specification with respect to the standard deviations of the disturbances $\varepsilon_{\tilde{y},t}$ and $\varepsilon_{\pi,t}$, allowing for a time variation in their volatilities $\sigma_{\tilde{y},t}$ and $\sigma_{\pi,t}$.

We then define r-star on the basis of the HLW specification. According to the neoclassical growth model, the steady-state real one-period interest rate is given by

$$r^* = \frac{1}{\sigma} g_c + \theta_p,$$

where σ represents the intertemporal elasticity of substitution in consumption, g_c denotes the growth rate of per capita consumption and θ_p is the rate of time preference. This relationship leads to standard monetary Dynamic Stochastic General Equilibrium (DSGE) models, where r^*, as given in Eq. (2.4), provides an effective method for determining the intercept in the applicable reduced-form interest rate rules (e.g., Taylor, 1993).

Given the above theoretical linkage and assuming a unit elasticity of substitution σ as in HLW, the r-star is then determined by

$$r_t^* = g_t + z_t, \tag{2.4}$$

where g_t is the trend growth rate of the natural rate of output and z_t captures other determinants of the natural rate of interest, such as households' rate of time preference.

Having defined the r-star, we finally introduce the system of transition equations governing the dynamics of the latent variables g_t and z_t and of the potential output y_t^* as

$$y_t^* = y_{t-1}^* + g_{t-1} + \varepsilon_{y^*,t} \qquad \varepsilon_{y^*,t} \sim \mathcal{N}(0, \sigma_{y^*,t}), \tag{2.5}$$

$$z_t = z_{t-1} + \varepsilon_{z,t} \qquad \varepsilon_{z,t} \sim \mathcal{N}(0, \sigma_z), \tag{2.6}$$

$$g_t = g_{t-1} + \varepsilon_{g,t} \qquad \varepsilon_{g,t} \sim \mathcal{N}(0, \sigma_g). \tag{2.7}$$

The transition of the potential output, y_t^*, follows a random walk, with a stochastic drift term given by the lagged trend growth rate, g_{t-1}. The variable z_t, which captures any persistent shocks to the r-star that are not captured by g_t, also follows a random walk similar to the trend growth rate. Again, this configuration is comparable to that in HLW, with one important exception being the volatility of the potential output, $\sigma_{y^*,t}$, which is assumed to be time-varying, thus completing the dynamic specification of the time-varying signal-to-noise ratio $\sigma_{y^*,t}^2/\sigma_{\tilde{y},t}^2$. As will be shown later, the signal-to-noise ratio is a central pillar in identifying persistent shocks in our setting. Accurate identification of its dynamics is expected to enhance the performance of the Kalman filter procedure used for inference on the latent state variables.

The parameters that are allowed to evolve in our modeling framework are collected in the time-varying parameter (TVP) vector

$$\lambda_t = (\sigma_{\tilde{y},t}, \sigma_{\pi,t}, \sigma_{y^*,t}, a_{\tilde{y},t}, b_{\pi,t})',$$

where

$$\sigma_{i,t} = \exp(\bar{\sigma}_{i,t}), \quad i = \{\tilde{y}, \pi, y^*\}, \quad a_{\tilde{y},t} = -\exp(\bar{a}_{\tilde{y},t}), \quad b_{\pi,t} = \exp\left(\bar{b}_{\pi,t}\right).$$

The remaining *static* parameters are collected in the vector

$$\theta_s = (a_1, a_2, b_1, \sigma_g, \sigma_z)',$$

where $\sigma_i = \exp(\bar{\sigma}_i)$, $i \in \{g, z\}$, and σ_i denotes the variance of the corresponding innovation.

To model the dynamics of the time-varying parameters, we adopt a score-driven approach that implies that λ_t is a function of the past information available at time $(t-1)$, that is, $\lambda_t = \lambda(\mathcal{I}_{t-1})$. The derivation of the score-driven recursions for modeling the dynamics of the elements of λ_t is addressed in the next Section 2.2.2.

2.2.2 The baseline score-driven conditionally Gaussian State-Space model

Due to their flexibility, *score-driven* models, such as the generalized autoregressive score (GAS) model (Creal et al., 2013) or, referring to the same concept, the dynamic conditional score model (DCS) (Harvey, 2013), have become a popular choice for introducing time-varying parameters in parametric time series models. GAS models can be viewed as a class of observation-driven models, where the conditional scaled score of the likelihood function drives the parameter dynamics. The

reliance on the full probability density conditional on past information, rather than on specific moments, makes the use of score-driven models particularly attractive.

Leveraging the feasibility of the approach in a multivariate Gaussian setting, Delle Monache et al. (2021) proposed the score-driven state-space model (SSM), where the time variation in the system matrices is modeled following a score-driven approach. An attractive reason for extending the standard linear SSM in a score-driven framework is that it provides a viable strategy for addressing dynamic parameter restrictions, which are often required for identification based on economic theory. In this respect, the score-driven SSM framework offers a flexible and tractable approach to make volatilities and NK structural relationships time-varying within the LW framework.

Before introducing the score-based updating mechanism for the time-varying parameters, it is helpful to first cast the structural framework described in Section 2.2.1 into a matrix state-space form. Specifically, the corresponding state-space model (Harvey, 1990) is expressed as follows.

$$\mathbf{y}_t = \boldsymbol{\Gamma}_t \mathbf{u}_t + \mathbf{Z}_t \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{H}_t) \tag{2.8}$$

$$\boldsymbol{\alpha}_t = \mathbf{T} \boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t \qquad \boldsymbol{\eta}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_t) \tag{2.9}$$

where $\mathbf{y}_t$ is the observation vector that collects contemporaneous endogenous variables, $\boldsymbol{\alpha}_t$ and $\mathbf{u}_t$ are the vectors of state and input variables, respectively, $\mathbf{T}$ is the state transition matrix, $\mathbf{Z}_t$ is the matrix of regression parameters linking $\mathbf{y}_t$ to the latent states, $\boldsymbol{\Gamma}_t$ plays the same role for the observed input covariates, $\mathbf{H}_t$ is the variance–covariance matrix of the observation error $\boldsymbol{\varepsilon}_t$, $\mathbf{Q}_t$ is the variance–covariance matrix of the system error $\boldsymbol{\eta}_t$. The output gap $\tilde{y}_t = y_t - y_t^*$ is expressed in percentage terms, consistent with the scaling of y_t and y_t^*. Formally, the model components and the system matrices are defined as follows.

$$\begin{aligned}
\mathbf{y}_t &= \begin{bmatrix} y_t, & \pi_t \end{bmatrix}', \\
\mathbf{u}_t &= \begin{bmatrix} y_{t-1}, & y_{t-2}, & r_{t-1}, & r_{t-2}, & \pi_{t-1}, & \pi_{t-2,4} \end{bmatrix}', \\
\boldsymbol{\alpha}_t &= \begin{bmatrix} y_t^*, & y_{t-1}^*, & y_{t-2}^*, & g_t, & g_{t-1}, & g_{t-2}, & z_t, & z_{t-1}, & z_{t-2} \end{bmatrix}',
\end{aligned}$$

$$\mathbf{Z}_t = \begin{bmatrix} 1 & -a_1 & -a_2 & 0 & -2a_{\tilde{y},t} & -2a_{\tilde{y},t} & 0 & -a_{\tilde{y},t}/2 & -a_{\tilde{y},t}/2 \\ 0 & -b_{\pi,t} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\boldsymbol{\Gamma}_t = \begin{bmatrix} a_1 & a_2 & a_{\tilde{y},t}/2 & a_{\tilde{y},t}/2 & 0 & 0 \\ b_{\pi,t} & 0 & 0 & 0 & b_1 & 1-b_1 \end{bmatrix}, \quad \mathbf{H}_t = \begin{bmatrix} \sigma_{\tilde{y},t} & 0 \\ 0 & \sigma_{\pi,t} \end{bmatrix},$$

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{Q}_t = \begin{bmatrix} \sigma_{y^*,t} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_g & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_z & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

It is worth noting again that, unlike HLW, in our framework, the system matrices $\mathbf{Z}_t$, $\mathbf{\Gamma}_t$, $\mathbf{H}_t$ and $\mathbf{Q}_t$ are allowed to vary over time. Furthermore, without affecting their stochastic nature, the assumed dynamic structure of the vector of time-varying parameters λ_t implies that these matrices are known conditionally on past information $\mathcal{I}_{t-1}$. Therefore, the resulting state-space model, similar to the one introduced by Delle Monache et al. (2021), is conditionally Gaussian, a property that allows us to construct the likelihood function in the usual prediction error decomposition form as in Eq (2.10) and use the standard Kalman filter (KF) recursions for optimal state estimation, thus retaining the simplicity of the classical linear Gaussian state space framework. Also, letting the system matrices $\mathbf{Z}_t$, $\mathbf{\Gamma}_t$, $\mathbf{H}_t$, and $\mathbf{Q}_t$ change over time, while still assuming that they are non-stochastic given the past, is a practical way to introduce nonlinearities into the state-space model. In fact, assuming that the parameters are determined by past observations, a feature that is compatible with the score-driven methodology, i) the system variables, observation, and state vectors are Gaussian, and ii) the observation and transition equations remain linear with respect to the state vector. However, these relationships are unconditionally nonlinear.

As in the case of time-invariant linear Gaussian state-space models, the model parameters θ are estimated by maximizing a Gaussian log-likelihood in its prediction error decomposition form, given by

$$\ell_t = \log p(y_t|\mathcal{I}_{t-1}, \theta) \propto -\frac{1}{2}\left(\log|F_t| + \nu_t' F_{t|t-1}^{-1}\nu_t\right), \tag{2.10}$$

where the variables involved in the computation of the likelihood function, namely the prediction error $\nu_{t|t-1}$ and its variance and covariance matrix F_t, along with the conditional mean of the state vector α_t and its mean squared error (MSE) matrix $P_{t|t}$, are estimated recursively in the KF. Formally, the *prediction* equations of the KF are given by

$$\begin{aligned}
\alpha_{t|t-1} &= T\alpha_{t-1|t-1} && (\alpha_t|\mathcal{I}_{t-1}) \sim \mathcal{N}(\alpha_{t|t-1}, P_{t|t-1}) \\
P_{t|t-1} &= TP_{t-1|t-1}T' + Q_t \\
y_{t|t-1} &= Z_t\alpha_{t|t-1} + \Gamma_t u_t && (y_t|\mathcal{I}_{t-1}) \sim \mathcal{N}(Z_t\alpha_{t|t-1} + \Gamma_t u_t, F_t) \\
\nu_{t|t-1} &= y_t - y_{t|t-1} \\
F_{t|t-1} &= Z_t P_{t|t-1} Z_t' + H_t
\end{aligned}$$

while the *filtering* equations are

$$\begin{aligned}
\alpha_{t|t} &= \alpha_{t|t-1} + K_t\nu_t \\
P_{t|t} &= P_{t|t-1} - K_t Z_t P_{t|t-1} = (I - K_t Z_t)P_{t|t-1} \\
K_t &= P_{t|t-1} Z_t' F_{t|t-1}^{-1}
\end{aligned}$$

where K_t is known as the *Kalman gain* matrix. As will be demonstrated, the score-driven model allows for high flexibility in updating the elements of the gain matrix, enhancing the ability to allocate the information content embedded in the prediction errors across the state variables. In contrast, in the static parameter model, these elements converge to a steady-state value with limited time variation, which hampers its efficiency in assigning appropriate weights to the state vector.

Constructing the model in a conditionally Gaussian form facilitates the use of KF to estimate the latent variables. However, it leaves the nontrivial challenge of jointly estimating both the state and the time-varying system matrices, a problem

addressed by Delle Monache et al. (2021) using a score-driven approach. The score-driven framework is readily incorporated into the conditional Gaussian state-space model by augmenting the KF recursions with a set of additional recursions that track the dynamics of the time-varying coefficients.

Following Creal et al. (2013) and Harvey (2013), the equation for the score-driven updating of the TVP vector λ_t takes the form

$$\lambda_{t+1} = \omega_\lambda + \mathbf{A}_\lambda \lambda_t + \mathbf{B}_\lambda s_t \quad , \quad s_t = S_{\lambda,t} u_{\lambda,t}, \tag{2.11}$$

where

$$\begin{aligned}
\omega_\lambda &= \left[\omega_{\lambda,\tilde{y}}, \quad \omega_{\lambda,\pi}, \quad \omega_{\lambda,y^*}, \quad 0, \quad 0\right]', \\
\mathbf{A}_\lambda &= \operatorname{diag}\left(\left[\mathrm{a}_{\lambda,\sigma_{\tilde{y}}}, \mathrm{a}_{\lambda,\sigma_\pi}, \mathrm{a}_{\lambda,\sigma_{y^*}}, \mathrm{a}_{\lambda,y^*}, \mathrm{a}_{\lambda,\pi}\right]\right), \\
\mathbf{B}_\lambda &= \operatorname{diag}\left(\left[\mathrm{b}_{\lambda,\sigma_{\tilde{y}}}, \mathrm{b}_{\lambda,\sigma_\pi}, \mathrm{b}_{\lambda,\sigma_{y^*}}, \mathrm{b}_{\lambda,y^*}, \mathrm{b}_{\lambda,\pi}\right]\right).
\end{aligned}$$

We impose diagonal parameter matrices $\mathbf{A}$ and $\mathbf{B}$ to remain fully consistent with the identification strategy employed in Laubach and Williams (2003). While allowing $\mathbf{B}$ to be full is feasible within a score-driven framework, doing so would introduce cross-equation volatility feedback, whereby shocks in one measurement equation affect the conditional variance of other measurement or state-transition disturbances. Such second-moment spillovers are not part of the original structural interpretation and would weaken the economic meaning of the mutually uncorrelated innovations assumption. The diagonal restriction ensures that innovations are not only contemporaneously uncorrelated but also dynamically separated in their volatility processes, preserving a clear mapping between structural shocks and their conditional variances. In addition, it improves parsimony and mitigates potential weak identification and numerical instability.

The vector s_t is the gradient of the likelihood function $u_{\lambda,t}$ scaled by the matrix square root of the inverse of the information matrix $S_{\lambda,t}$[1]

$$u_{\lambda,t} = \frac{\partial \log p(\mathbf{y_t} \mid \lambda_t; \theta)}{\partial \lambda_t} \quad , \quad S_{\lambda,t} = -\mathbf{E_t}\left(\frac{\partial^2 \ell_t}{\partial f_t \partial f_t'}\right)^{-\frac{1}{2}} = \mathcal{I}_t^{-\frac{1}{2}}.$$

The gradient and the information matrix are computed analytically in parallel with the KF recursion as

$$u_{\lambda,t} = \frac{1}{2}\dot{F}_t (F_t \otimes F_t)^{-1} \mathrm{vec}(\nu_t \nu_t' - F_t) - 2\dot{V}_t F_t^{-1} \nu_t, \tag{2.12}$$

$$\mathcal{I}_t = \frac{1}{2}\dot{F}_t (F_t \otimes F_t)^{-1} \dot{F}_t + 2\dot{V}_t F_t^{-1} \dot{V}_t, \tag{2.13}$$

where $\dot{V}_t = \partial \nu_t / \partial \lambda_t$ and $\dot{F}_t = \partial \mathrm{vec}(F_t) / \partial \lambda_t$ measure the sensitivity of the prediction error ν_t and its variance F_t with respect to λ_t. The Jacobian matrices involved in the computation of $\dot{V}_t$ and $\dot{F}_t$ are derived following the approach in Delle Monache et al. (2021). However, it should be noted that, compared to their specification, the observation equation in our model includes an additional regression term, $\mathbf{\Gamma}_t \mathbf{u}_t$. This modification affects the derivation of $\dot{V}_t$, which must be adjusted accordingly to account for the contribution of $\mathbf{\Gamma}_t$.

[1] To improve numerical stability, we replace $S_{\lambda,t}$ with its smoothed version $\tilde{S}_{\lambda,t} = (1-\kappa) S_{\lambda,t} + \kappa \tilde{S}_{\lambda,t-1}$, where the smoothing parameter κ is estimated jointly with the other model parameters through ML.

The nonlinear mappings from the TVP vector λ_t to the system matrices, and the associated Jacobians employed by the score-driven filter, are as follows.

Time variation in the input matrix Γ evolves through the equation

$$vec(\Gamma_t) = S_{0,\Gamma} + S_{1,\Gamma}\psi(S_{2,\Gamma}\lambda_t), \tag{2.14}$$

where the components take the following forms:

$$vec(\Gamma_t) = \begin{bmatrix} a_1 \\ b_{\pi,t} \\ a_2 \\ 0 \\ a_{\tilde{y},t}/2 \\ 0 \\ a_{\tilde{y},t}/2 \\ 0 \\ 0 \\ b_1 \\ 0 \\ (1-b_1) \end{bmatrix}, \qquad S_{0,\Gamma} = \begin{bmatrix} a_1 \\ 0 \\ a_2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ b_1 \\ 0 \\ (1-b_1) \end{bmatrix}, \qquad S_{1,\Gamma} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\psi_\Gamma = \begin{bmatrix} a_{\tilde{y},t}/2 \\ b_{\pi,t} \\ a_{\tilde{y},t}/2 \end{bmatrix}, \qquad S_{2,\Gamma} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The related Jacobian matrix is

$$\dot{\Gamma}_t = S_{1,\Gamma}\dot{\Psi}_{\Gamma,t}S_{2,\Gamma}, \tag{2.15}$$

where

$$\dot{\Psi}_{\Gamma,t} = \frac{\partial\psi(S_{2,\Gamma}\lambda_t)}{\partial\lambda_t'} = \frac{\partial\psi_{\Gamma,t}(S_{2,\Gamma}\lambda_t)}{\partial(S_{2,\Gamma}\lambda_t)'}\frac{\partial(S_{2,\Gamma}\lambda_t)}{\partial\lambda_t'} = \begin{bmatrix} a_{\tilde{y},t}/2 & 0 \\ 0 & b_{\pi,t} \\ a_{\tilde{y},t}/2 & 0 \end{bmatrix}.$$

Time variation in the system matrix Z evolves through the equation

$$vec(Z_t) = S_{0,Z} + S_{1,Z}\psi(S_{2,Z}\lambda_t), \tag{2.16}$$

where

$$vec(Z_t) = \begin{bmatrix} 1 \\ 0 \\ -a_1 \\ -b_{\pi,t} \\ -a_2 \\ 0 \\ 0 \\ 0 \\ -a_{\tilde{y},t} \cdot 2 \\ 0 \\ -a_{\tilde{y},t} \cdot 2 \\ 0 \\ 0 \\ 0 \\ -a_{\tilde{y},t}/2 \\ 0 \\ -a_{\tilde{y},t}/2 \\ 0 \end{bmatrix}, \qquad S_{Z,0} = \begin{bmatrix} 1 \\ 0 \\ -a_1 \\ 0 \\ -a_2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad S_{Z,1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\psi_Z = \begin{bmatrix} -a_{\tilde{y},t} \cdot 2 \\ -b_{\pi,t} \\ -a_{\tilde{y},t} \cdot 2 \\ -a_{\tilde{y},t}/2 \\ -a_{\tilde{y},t}/2 \end{bmatrix}, \qquad S_{Z,2} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The associated Jacobian matrix is

$$\dot{Z}_t = S_{1,Z} \dot{\Psi}_{Z,t} S_{2,Z}, \tag{2.17}$$

where

$$\dot{\Psi}_{Z,t} = \frac{\partial \psi(S_{2,Z}\lambda_t)}{\partial \lambda_t'} = \frac{\partial \psi_{Z,t}(S_{2,Z}\lambda_t)}{\partial (S_{2,Z}\lambda_t)'} \frac{\partial (S_{2,Z}\lambda_t)}{\partial \lambda_t'} = \begin{bmatrix} -a_{\tilde{y},t} \cdot 2 & 0 \\ 0 & -b_{\pi,t} \\ -a_{\tilde{y},t} \cdot 2 & 0 \\ -a_{\tilde{y},t}/2 & 0 \\ -a_{\tilde{y},t}/2 & 0 \end{bmatrix}.$$

Time variation in the system matrix H is defined as

$$vec(H_t) = S_{0,H} + S_{1,H} \psi_H(S_{2,H}\lambda_t), \tag{2.18}$$

where

$$vec(H_t) = \begin{bmatrix} \sigma_{\tilde{y},t} \\ 0 \\ 0 \\ \sigma_{\pi,t} \end{bmatrix}, \qquad S_{0,H} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad S_{1,H} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\psi_H = \begin{bmatrix} \sigma_{\tilde{y},t} \\ \sigma_{\pi,t} \end{bmatrix}, \qquad S_{2,H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

The Jacobian matrix is

$$\dot{H}_t = S_{1,H}\dot{\Psi}_{H,t}S_{2,H},$$

where

$$\dot{\Psi}_{H,t} = \frac{\partial\psi(S_{2,H}\lambda_t)}{\partial\lambda_t'} = \frac{\partial\psi_{H,t}(S_{2,H}\lambda_t)}{\partial(S_{2,H}\lambda_t)'}\frac{\partial(S_{2,H}\lambda_t)}{\partial\lambda_t'} = \begin{bmatrix}\sigma_{\tilde{y},t} & 0\\ 0 & \sigma_{\pi,t}\end{bmatrix}.$$

Time variation in the system matrix Q follows as

$$vec(Q_t) = S_{0,Q} + S_{1,Q}\psi_Q(S_{2,Q}\lambda_t), \tag{2.19}$$

where

$$vec(Q_t) = \begin{bmatrix}\sigma_{y^*,t}\\ 0_{(29\times1)}\\ \sigma_g\\ 0_{(29\times1)}\\ \sigma_z\\ 0_{(20\times1)}\end{bmatrix}, \qquad S_{0,Q} = \begin{bmatrix}0_{(30\times1)}\\ \sigma_g\\ 0_{(29\times1)}\\ \sigma_z\\ 0_{(20\times1)}\end{bmatrix}, \qquad S_{1,Q} = \begin{bmatrix}1\\ 0_{(80\times1)}\end{bmatrix},$$

$$\psi_Q = \begin{bmatrix}\sigma_{y^*,t}\end{bmatrix}, \qquad S_{2,Q} = \begin{bmatrix}0 & 0 & 1 & 0 & 0\end{bmatrix}.$$

The associated Jacobian matrix is

$$\dot{Q}_t = S_{1,Q}\dot{\Psi}_{Q,t}S_{2,Q}, \tag{2.20}$$

where

$$\dot{\Psi}_{Q,t} = \frac{\partial\psi(S_{2,Q}\lambda_t)}{\partial\lambda_t'} = \frac{\partial\psi_{Q,t}(S_{2,Q}\lambda_t)}{\partial(S_{2,Q}\lambda_t)'}\frac{\partial(S_{2,Q}\lambda_t)}{\partial\lambda_t'} = \sigma_{y^*,t}.$$

Lastly, the Jacobian associated with the prediction errors takes the form:

$$\dot{V}_t = \frac{\partial v_t}{\partial\lambda_t'} = \left[\frac{\partial v_t}{\partial \mathrm{vec}(Z_t)'}\frac{\partial \mathrm{vec}(Z_t)}{\partial\lambda_t'} + \frac{\partial v_t}{\partial\alpha_t'}\frac{\partial\alpha_t}{\partial\lambda_t'} + \frac{\partial v_t}{\partial \mathrm{vec}(\Gamma_t)'}\frac{\partial \mathrm{vec}(\Gamma_t)}{\partial\lambda_t'}\right]$$

$$= -\left[(\boldsymbol{\alpha}_t'\otimes I_n)\frac{\partial\,\mathrm{vec}(Z_t)}{\partial\boldsymbol{\lambda}_t'} + Z_t\otimes\frac{\partial\boldsymbol{\alpha}_t}{\partial\boldsymbol{\lambda}_t'} + (\boldsymbol{u}_t'\otimes I_n)\frac{\partial\,\mathrm{vec}(\Gamma_t)}{\partial\boldsymbol{\lambda}_t'}\right]$$

$$= -\left[(\boldsymbol{\alpha}_t'\otimes I_n)\dot{Z}_t + (\boldsymbol{u}_t'\otimes I_n)\dot{\Gamma}_t\right], n = 2.$$

In our implementation, to limit the proliferation of model parameters, we have assumed that the matrices $\mathbf{A}_\lambda$ and $\mathbf{B}_\lambda$ are diagonal.

The diagonal elements of the matrix $\mathbf{B}_\lambda$ ($\mathrm{b}_{\lambda,i}$) determine how sensitive the TVP vector is to the score s_t. Regarding the matrix $\mathbf{A}_\lambda$, its diagonal elements ($\mathrm{a}_{\lambda,i}$) contribute to the degree of persistence of the time-varying parameters by measuring the speed at which the parameter reverts to its long-run level in a shock-free framework. In other words, the diagonal elements of $\mathbf{A}_\lambda$ determine the inertia of the information embedded in the prediction error up to time $t-1$. As

discussed in the following section, the above interpretation of the two weighting matrices plays a crucial role in the context of the present study.

Also, since the time variation of the slope parameters is assumed to follow a random walk in the IS and Phillips curve equations, we have imposed $\omega_{\lambda,i} = 0$ and $\mathrm{a}_{\lambda,i} = 1$ in the updating equations for $a_{\tilde{y},t}, b_{\pi,t}$. The vector of static parameters driving the dynamics of the TVP vector λ_t is then given by

$$\boldsymbol{\theta}_\lambda = \{\omega_{\lambda,i}, \mathbf{A}_{\lambda,i}, \mathbf{B}_{\lambda,i}\}, \quad i = \{\sigma_{\tilde{y}}, \sigma_\pi, \sigma_{y^*}, a_{\tilde{y}}, b_\pi\}.$$

The total vector of model parameters to be estimated maximizing the log-likelihood in Eq. (2.10) is $\boldsymbol{\theta} = (\boldsymbol{\theta}_s', \boldsymbol{\theta}_\lambda')'$, which includes the previously introduced vector of static parameters $\boldsymbol{\theta}_s$ and the vector of TVP-related coefficients $\boldsymbol{\theta}_\lambda$.

2.2.3 An augmented accelerating score-driven framework for TVP updating

The GAS framework allows for score-driven TVPs but relies on the assumption that the dynamic law governing their evolution is time-invariant. However, changing economic conditions and state dependencies, which typically characterize many economic phenomena, can cause the features of the variation pattern to evolve. Furthermore, the flexibility of GAS models is limited by the assumption that the TVPs are linearly dependent on past score values. Blasques et al. (2019) recognized and addressed such difficulties and proposed the accelerating GAS (aGAS) model, which assigns a time-varying weight to the score in the parameter updating equation.

The aGAS model includes an additional updating equation in which the first-order autocorrelation of the conditional scores determines the weight assigned to the score in updating the dynamic parameters. When recent score innovations have the same sign (positive first-order autocorrelation), the adjustment of the current TVP needs to accelerate faster than in a period when these innovations have mixed signs. Accordingly, the updating procedure will accelerate when a set of consecutive innovations has the same signs. As a result, the aGAS provides an intuitive way to make the TVP adjustment speed adaptive, thus improving the local fit of the model. In their analysis, Blasques et al. (2019) find that this mechanism is well suited to capture sudden volatility shifts, adapting the volatility dynamics much faster than in the standard GARCH model. In addition, they find that in the presence of outliers, the application of the aGAS approach in GARCH-type models can mitigate the adverse effects of tail events by attenuating the impact of outliers faster.

Considering the aforementioned aspects in this study, the large transitory shocks experienced during the COVID-19 pandemic serve as a striking example of rapid changes in volatility dynamics. As Holston et al. (2023) report, the standardized auxiliary residuals of the output gap display severe divergences from their pre-pandemic level of standard deviation due to the economic shutdowns and reopenings caused by the exogenous shocks. Consistent with the procedure for testing the presence of outliers using auxiliary residuals proposed by Harvey and Koopman (1992) and Koopman et al. (1998), this observation confirms that the estimated output gap has been heavily affected by extreme observations associated with the pandemic.

Since the main objective in estimating the r-star is to properly disentangle transitory disturbances from permanent or highly persistent shocks using the KF, it

is crucial to model volatility persistence adequately. Volatility directly influences the elements of the Kalman gain matrix, which determines the weight assigned to the information content of the prediction errors in the state estimation process. As suggested by time series evidence, highly persistent changes in real GDP growth, labor productivity growth, and the real interest rate are prone to be mistaken for volatile transitory shocks (Holston et al., 2017). Therefore, maximum likelihood (ML) estimates of the standard deviations in the observation equation are potentially contaminated by persistent changes. As a result, the standard deviations of the state innovations are likely to be biased towards zero when estimated by maximum likelihood, which is the so-called *pile-up* problem discussed in Stock (1994).

To overcome these problems, we propose implementing an aGAS model, which provides a flexible approach to identifying transitory shocks through adaptive modeling of volatility persistence. As mentioned above, this task is achieved by assigning a time-varying weight to the conditional score. Furthermore, to improve the flexibility of the aGAS modeling approach, we consider a variant of the model initially proposed by Blasques et al. (2019) where, in the updating mechanism, the autoregressive coefficient associated with the lagged volatility is also allowed to vary over time. Since this coefficient determines the extent to which current volatility depends on its past values and thus governs its inertia, it plays a prominent role in determining the persistence of volatility clustering over time. This modified version of accelerating GAS will be referred to as augmented accelerating GAS (aaGAS).

In the aaGAS model, the standard score-driven recursion to update the TVP vector takes the form

$$\lambda_{t+1} = \omega_\lambda + \mathbf{A}_{\lambda,t}\lambda_t + \mathbf{B}_{\lambda,t}s_t \quad , \quad s_t = S_{\lambda,t}u_{\lambda,t}, \tag{2.21}$$

where the persistence parameters, given by the diagonal elements of $\mathbf{A}_{\lambda,t}$ ($\mathbf{a}_{\lambda,t}$) and $\mathbf{B}_{\lambda,t}$ ($\mathbf{b}_{\lambda,t}$), become time-varying. The aGAS model is obtained as a special case setting $\mathbf{A}_{\lambda,t} = \mathbf{A}_\lambda$. In our framework, considering the extreme shifts in the volatility of the output gap disturbances during the COVID-19 pandemic, we have decided to enhance the baseline score-driven model presented in Section 2.2.2 by introducing an aaGAS update only for the standard deviation of the observation noise in the IS curve equation. To keep the model parsimonious and therefore confine the dimension of the TVP-related static parameter vector, the weight parameters associated with the other elements of the TVP vector remain time-invariant, leading to the following modified configuration for the $\mathbf{A}_{\lambda,t}$ and $\mathbf{B}_{\lambda,t}$ matrices present in Eq. (2.21).

$$\mathbf{A}_{\lambda,t} = \operatorname{diag}\left(\left[\boxed{\mathrm{a}_{\lambda,\sigma_{\tilde{y}},t}},\ \mathrm{a}_{\lambda,\sigma_\pi},\ \mathrm{a}_{\lambda,\sigma_{y^*}},\ \mathrm{a}_{\lambda,y^*},\ \mathrm{a}_{\lambda,\pi}\right]\right),$$
$$\mathbf{B}_{\lambda,t} = \operatorname{diag}\left(\left[\boxed{\mathrm{b}_{\lambda,\sigma_{\tilde{y}},t}},\ \mathrm{b}_{\lambda,\sigma_\pi},\ \mathrm{b}_{\lambda,\sigma_{y^*}},\ \mathrm{b}_{\lambda,y^*},\ \mathrm{b}_{\lambda,\pi}\right]\right),$$

where the variation in $\mathrm{a}_{\lambda,\sigma_{\tilde{y}},t}$ and $\mathrm{b}_{\lambda,\sigma_{\tilde{y}},t}$ (boxed) is modeled by the aaGAS specification. For simplicity, hereafter we denote these as $\mathrm{a}_{\lambda,t}$ and $\mathrm{b}_{\lambda,t}$, respectively.

Following Blasques et al. (2019), the weighting of the ith component of the score innovation is achieved as follows.

$$b_{\lambda,t} = j(f_{t+1};\theta) \quad , \quad f_{t+1} = \omega_f + \mathrm{a}_f f_t + \mathrm{b}_f s_{f,t}, \tag{2.22}$$

where $j(\cdot)$ denotes a logistic function that characterizes how the additional updating equation governs the time variation in $\mathrm{b}_{\lambda,t}$. The time-varying f_{t+1} follows a first-order autoregressive process with intercept ω_f, AR coefficient a_f, weight parameter b_f, and innovation term $s_{f,t}$, which is given by

$$s_{f,t} = C_{f,t} u_{f,t} \quad , \quad u_{f,t} = \frac{\partial \log p(y_t \mid \lambda_t; \theta)}{\partial f_t} \propto u_{\lambda,t} u_{\lambda,t-1}, \tag{2.23}$$

where $C_{f,t}$ is a function of the scaling factors $S_{\lambda,t}$ and $S_{\lambda,t-1}$ introduced earlier.

More precisely, since the aaGAS update is applied exclusively to the first element of the TVP vector, we consider only the first diagonal element of the scaling matrices $S_{\lambda,t}$ and $S_{\lambda,t-1}$, such that:

$$C_{f,t} = [S_{\lambda,t}]_{11} \cdot [S_{\lambda,t-1}]_{11},$$

where

$$[S_{\lambda,t}]_{11} = ([\mathcal{I}_t]_{11})^{-\frac{1}{2}}$$

denotes the square root of the inverse of the first diagonal element of the time t information matrix. This scalar scaling term ensures that only the uncertainty related to the output gap volatility enters the aaGAS update so that the current and lagged scores are scaled by their corresponding marginal Fisher information.

To complete the local dependence model of volatility persistence in Eq. (2.21), we introduce $\mathrm{a}_{\lambda,t}$ given as

$$\mathrm{a}_{\lambda,t} = k(g_{t+1}; \theta) \quad , \quad g_{t+1} = \omega_g + \mathrm{a}_g g_t + \mathrm{b}_g s_{g,t} \quad , \quad s_{g,t} = [S_{\lambda,t}]_{11} \cdot u_{g,t}, \tag{2.24}$$

where the link function and the structure of the updating equation are analogous to Eq. (2.22), with the innovation term $u_{g,t}$.

Assuming $\partial \lambda_{t-1} / \partial g_t = 0$, we obtain

$$\begin{aligned} \frac{\partial \log p}{\partial g_t} &= \frac{\partial \log p}{\partial \lambda_t} \frac{\partial \lambda_t}{\partial g_t} \\ &= u_{\lambda,t} \frac{\partial \lambda_t}{\partial g_t} \\ &= u_{\lambda,t} \left(\mathrm{b}_{\lambda,t-1} \frac{\partial \lambda_{t-1}}{\partial g_t} + \mathrm{a}_{\lambda,t} \frac{\partial s_{\lambda,t-1}}{\partial g_t} \right) \\ &\propto u_{\lambda,t} \lambda_{t-1}. \end{aligned} \tag{2.25}$$

Thus,

$$u_{g,t} = \frac{\partial \log p(y_t \mid \lambda_t; \theta)}{\partial g_t} \propto u_{\lambda,t} \lambda_{t-1}. \tag{2.26}$$

Therefore, Eq. (2.26) reveals that when the weighting coefficient in the score-driven updating equation becomes time-varying, it is proportional to the product of the current conditional score and the lagged realization of the time-varying parameter itself, thus relating to the first-order cross-correlation between these variables. One may interpret this result within the present context as follows: a sharp change in the conditional score, materialized by a large shock such as an event associated with the COVID-19 pandemic, implies a sudden downward shift in volatility persistence. Therefore, the model is expected to identify large transitory shocks more effectively by assigning less weight to the AR coefficient in the parameter updating equation.

2.3 Empirical analysis

This section presents the results of an empirical application of the models introduced in Section 2.2 to US data. First, Sections 2.3.1 and 2.3.2 provide brief descriptions of the dataset used for the analysis and the estimation strategy, respectively, while Section 2.3.3 presents the estimation results, Section 2.3.4 focuses on the estimated time-varying volatilities, the estimated dynamic IS and Phillips curves are presented in Section 2.3.5, and finally Section 2.3.6 reports and discusses the empirical results on the estimation of the r-star and the output gap.

2.3.1 The data

The dataset used in this study is the same as that of Holston (2017), but has been extended to cover the period up to 2024Q4. Therefore, for a brief overview, we report only the summary of the variables in Table 2.2. A detailed description of the variables can be found in the Data Appendix of HLW. More details and the updated quarterly database are available on the website of the Federal Reserve Bank of New York (`https://www.newyorkfed.org/research/policy/rstar`).

Table 2.2: Variable Definitions

Variable	**Description**
Inflation	Annualized quarterly growth rate of core PCE inflation
Inflation expectations	Four quarter moving average of past inflation
Short-term interest rate	Annualized nominal federal funds rate
Output	Logarithm of real GDP

2.3.2 Model specification and estimation

We estimate three variants of the baseline time-varying parameter (TVP) conditionally Gaussian state-space model, each employing a distinct score-driven updating mechanism for the dynamics of the TVP vector. First, we estimate the model using the standard GAS specification. Next, we consider a version with an accelerating score-driven update (aGAS). Finally, we estimate the full specification, which incorporates an augmented accelerating scheme (aaGAS) for the volatility of the output gap innovations. The model log-likelihood in Eq. (2.10) is evaluated by the KF, augmented with the score-driven recursions needed to update the time-varying parameters present in the system matrices. Estimation of the vector of model parameters θ, including those kept time-invariant within the system matrices (θ_s) and the TVP-related ones (θ_λ), is carried out by numerically maximizing the log-likelihood function with respect to $\theta = (\theta_s, \theta_\lambda)$. The confidence intervals for the elements of the TVP vector λ_t and the state estimates are constructed by integrating the simulation-based inferential techniques of Blasques et al. (2016) and Hamilton (1986) to account for both parameter and filter uncertainty. More specifically, in addition to generating parameter vectors from a multivariate normal distribution, with mean and variance-covariance matrix, respectively, given by the ML estimate of the static parameter vector $\hat{\theta}$ and its robust variance-covariance matrix estimate, we simultaneously draw model-based random trajectories of the state vector. As prior for the initial state, we assume a multivariate normal distribution with the mean and variance-covariance matrix given by the smoothed estimates

of the initial state and its variance-covariance matrix, respectively. Therefore, the confidence intervals constructed from the simulated sequences of parameter and state vectors obtained using the Kalman filter with the TVP updating algorithm capture both filtering and parameter estimation uncertainty. In addition, to assess the overall precision of the state variable estimates and facilitate comparison with the HLW model, we compute their associated standard errors following the methodology introduced by Hamilton (1986). In what follows, we report standard errors and 68% confidence intervals based on 5,000 simulations.

2.3.3 Parameter estimates

Before turning to the visual representation of the estimated dynamic parameters and state variables, Table 2.3 presents a first summary of the estimation results. In addition to the score-driven models, the table also reports the estimation results for the COVID-adjusted version of the benchmark HLW model discussed by Holston et al. (2023). All models consider the period from 1960Q1 to 2024Q4, covering the complete data set available at the time of writing.

The top panel of the table reports the parameter estimates for the static parameters in θ_s, $(a_1, a_2, b_1, \sigma_g, \sigma_z)$, along with their robust standard errors. For both score-driven specifications, the estimated values are broadly in line with the HLW estimates, except for the standard deviations of the trend growth rate and the innovations of the z-factor, σ_g and σ_z, which are estimated to be lower under the score-driven models. At this point, it is worth noting that we slightly depart from the related HLW procedure, where the Stock and Watson (1998) median unbiased estimator (MUE) is used due to the pile-up problem discussed in Section 2.2.3. Our choice aligns with Buncic (2024), who estimates σ_g and σ_z with ML and finds little evidence for the pile-up problem. As the estimated parameters suggest, this finding is also supported by our ML estimates, likely due to the improved identification of transitory shocks facilitated by time-varying volatilities. Concerning the estimation of the TVP vector, in order to enable the comparison with HLW, we summarize the elements of λ_t in terms of their sample averages in the bottom panel of the table. Specifically, we find that the mean values of $\sigma_{\tilde{y}}$, σ_π, and σ_{y^*} are consistent with the corresponding HLW estimates. Regarding the two slope parameters, $a_{\tilde{y}}$ and b_π, the first, associated with the IS curve, is well aligned with the HLW estimate. In contrast, the sample average of the latter, corresponding to the Phillips curve, is half the value reported by HLW when using the GAS specifications.[2]

Turning to the assessment of uncertainty in the state variables, the lower panel of Table 2.3 compares the standard errors across the models under consideration. Focusing first on potential output, the sample-average standard errors in the GAS models are higher than in the HLW model, indicating somewhat greater estimation uncertainty on average. However, this difference is only marginal relative to the aGAS and aaGAS models. Considering the final observation, the standard GAS model exhibits a similar gap, while both the aGAS and aaGAS models produce nearly identical standard errors. This suggests that the full-sample divergence from the HLW model in potential output estimation uncertainty is not driven by the onset of the COVID-19 pandemic.[3] Most importantly, the sample average

[2]The slope parameter of the Phillips curve estimated under the original LW configuration is closer to the sample average of our estimates, i.e., 0.047 over the same estimation interval.

[3]We find that the standard errors of potential output generated by the aaGAS model are, on average, similar to those obtained from the original Laubach and Williams (2003) model

Table 2.3: Model estimation results

Parameter	**HLW**	**standard GAS**	**aGAS**	**aaGAS**
		Static [θ_s]		
a_1	1.417	1.514	1.462	1.544
	[0.103]	[0.097]	[0.172]	[0.027]
a_2	-0.483	-0.545	-0.503	-0.578
	[0.105]	[0.084]	[0.233]	[0.035]
b_1	0.689	0.617	0.618	0.618
	[0.041]	[0.079]	[0.093]	[0.074]
σ_g	0.137	0.128	0.080	0.080
		[0.005]	[0.003]	[0.001]
σ_z	0.113	0.056	0.059	0.045
		[0.022]	[0.061]	[0.001]
$\sigma_{r^*} = \sqrt{\sigma_g^2 + \sigma_z^2}$	0.188	0.139	0.099	0.092
		Time Varying [λ_t]		
$a_{\tilde{y}}$	-0.068	-0.054	-0.063	-0.057
	[0.017]			
b_π	0.080	0.040	0.041	0.040
	[0.026]			
$\sigma_{\tilde{y}}$	0.437	0.386	0.460	0.391
	[0.094]			
σ_π	0.791	0.753	0.743	0.739
	[0.027]			
σ_{y^*}	0.503	0.511	0.474	0.515
	[0.077]			
		S.E. (sample averages)		
r^*	1.207	0.760	0.751	0.700
g	0.407	0.400	0.310	0.309
y^*	1.520	2.023	1.778	1.809
z	1.137	0.644	0.683	0.627
		S.E. (final observations)		
r^*	1.595	1.073	1.038	0.910
g	0.583	0.642	0.488	0.452
y^*	2.090	2.878	2.089	2.084
z	1.484	0.860	0.916	0.789
Log-Likelihood	-600.70	-570.44	-568.43	-565.81
AIC	1235.4	1174.89	1174.87	1173.62

Note: Robust standard errors (S.E.) are reported in parentheses. σ_g is expressed at an annual rate. The values for the time-varying counterparts of $\sigma_{\tilde{y}}$, σ_π, σ_{y^*}, $a_{\tilde{y}}$ and b_π under the GAS specifications correspond to their sample averages. Standard errors for σ_g and σ_z under the HLW model are not reported, as these are implied parameters. Specifically, they are computed as $\sigma_g = \lambda_g \sigma_{y^*}$ and $\sigma_z = \frac{\lambda_z \sigma_{\tilde{y}}}{a_{\tilde{y}}}$, where λ_g and λ_z are obtained from a preliminary ML estimation.

standard error of the r-star in the aaGAS model is 0.7 percentage points, which is nearly half the magnitude observed in the HLW model, highlighting a substantial reduction in uncertainty of the natural rate estimate when the full information set is used. A comparable gain in precision is also evident in the standard errors of the final observation, reflecting lower uncertainty in the real-time state estimates of the score-driven models. It should be noted that the random walk component, z, accounts for the largest improvement in the precision of r-star estimates and is the

configuration with the COVID-related modification introduced by Holston et al. (2023), and lower for the final observation, which reports 1.8 and 2.5 percentage points, respectively.

primary contributor to the estimation uncertainty of the natural rate of interest reported under all models, as well as for both the sample average and the final observation estimates.[4]

Overall, the log-likelihood and information criteria indicate that the score-driven framework significantly improves model fit, with the aaGAS specification slightly outperforming the other GAS variants.

Table 2.4: TVP Related Parameter Estimates

Parameter	**standard GAS**	**aGAS**	**aaGAS**	**Parameter**	**aGAS**	**aaGAS**
$a_{\lambda,\sigma_{\tilde{y}}}$	0.826	0.815	0.463	$a_{f,\sigma_{\tilde{y}}}$	0.713	0.837
	[0.053]	[0.002]			[0.652]	[0.016]
$a_{\lambda,\sigma_{\pi}}$	0.900	0.896	0.898	$b_{f,\sigma_{\tilde{y}}}$	0.007	0.006
	[0.050]	[0.048]	[0.042]		[0.054]	[0.002]
$a_{\lambda,\sigma_{y^*}}$	0.838	0.858	0.740	$\omega_{f,\sigma_{\tilde{y}}}$	0.450	0.490
	[0.145]	[0.006]	[0.009]		[0.348]	[0.005]
$b_{\lambda,\sigma_{\tilde{y}}}$	0.498	0.328	0.440	$a_{g,\sigma_{\tilde{y}}}$		0.845
	[0.268]					[0.026]
$b_{\lambda,\sigma_{\pi}}$	0.142	0.140	0.139	$b_{g,\sigma_{\tilde{y}}}$		0.030
	[0.032]	[0.030]	[0.031]			[0.001]
$b_{\lambda,\sigma_{y^*}}$	0.275	0.305	0.221	$\omega_{g,\sigma_{\tilde{y}}}$		0.464
	[0.096]	[0.185]	[0.011]			[0.074]
$b_{\lambda,\tilde{y}}$	0.118	0.074	0.078			
	[0.116]	[0.230]	[0.016]			
$b_{\lambda,\pi}$	0.006	0.005	0.003			
	[0.008]	[0.014]	[0.004]			
$\omega_{\lambda,\sigma_{\tilde{y}}}$	0.719	0.750	0.366			
	[0.345]	[0.108]	[0.047]			
$\omega_{\lambda,\sigma_{\pi}}$	0.945	0.940	0.940			
	[0.036]	[0.040]	[0.027]			
$\omega_{\lambda,\sigma_{y^*}}$	0.804	0.809	0.708			
	[0.368]	[0.199]	[0.013]			
κ	0.014	0.018	0.036			
	[0.007]	[0.014]	[0.004]			

Note: Robust standard errors are reported in parentheses. The values for the time-varying counterpart of $a_{\lambda,\sigma_{\tilde{y}}}$ under aGAS and $a_{\lambda,\sigma_{\tilde{y}}}$, $b_{\lambda,\sigma_{\tilde{y}}}$ under the aaGAS model correspond to their sample averages.

The additional static parameters that govern the score-driven updates, collected in the vector θ_λ, are reported in Table 2.4. Most parameters are significant at the 10% level in both score-driven models. The most notable exception is given by the score weight associated with the standard deviation of the output gap, $b_{\lambda,\sigma_{\tilde{y}}}$, indicating some uncertainty in identifying the dynamics of the underlying coefficient in the standard and accelerating GAS models. As $a_{\lambda,\sigma_{\tilde{y}}}$ and $b_{\lambda,\sigma_{\tilde{y}}}$ are modeled as time-varying in the aaGAS framework, we report the time averages of the estimated coefficients to allow comparison with the level of the corresponding GAS estimates.

[4]We note that the gains in the estimation precision are similar or even greater when comparing the standard errors of the r-star and its components obtained by our models with those by the Laubach and Williams (2003) specification.

2.3.4 Time-varying volatilities

In this section, we compare the estimates of time-varying volatilities obtained by the considered models.

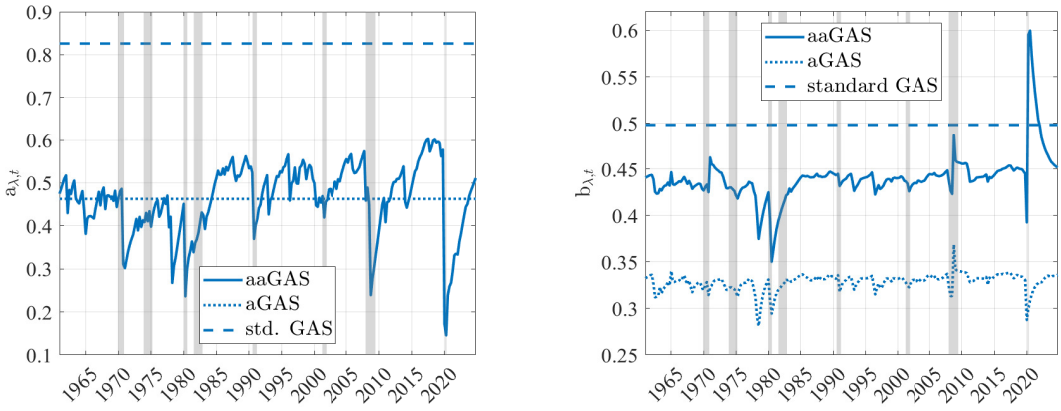

Figure 2.1: Dynamic persistence parameters of the output gap noise volatility estimated by the aaGAS (solid line), aGAS (dotted line) and standard GAS (dashed line) specifications. The left-hand side panel displays the inertia coefficient $a_{\lambda,t}$, the right-hand side panel shows the weight of the score innovations $b_{\lambda,t}$.

First, Figure 2.1 reports the evolution of the parameters governing the state-dependent persistence of output gap volatility. The time-varying AR coefficient $a_{\lambda,t}$ in the aaGAS model exhibits substantial variation over the sample period. Notably, its dynamics appear to be correlated with those of output, as sharp declines in the coefficient coincide with US recessionary periods characterized by pronounced contractions in output. In general, $a_{\lambda,t}$ decreases with the magnitude of output shocks. These declines are typically followed by a gradual recovery as economic disruptions dissipate. This behavior is particularly pronounced during the COVID-19 pandemic, reflecting highly volatile economic activity.

Over the same period, the parameter $b_{\lambda,t}$ in the aaGAS model, which measures the weight on score innovations, exhibits a sharp increase following an initial decrease. This variation enables a faster adjustment to the new level of volatility. In contrast, the corresponding parameter in the aGAS model shows a gradual reversion to its long-term reference value, resulting in a more sluggish adjustment. Consequently, the faster adjustment associated with the lower level of $a_{\lambda,t}$ in the aaGAS model implies that the overall persistence of the output gap during the COVID period is predominantly driven by score innovations rather than by the inertial component of its dynamics.

Next, Figures 2.2 and 2.3 show the time plots of the output gap and potential output volatilities, respectively, comparing the time-varying volatility dynamics returned by the score-driven models with the volatility level fitted by the COVID-adjusted HLW model. The latter assumes a constant volatility level, except for an adjustment restricted to the COVID-19 period.

For the output gap, as shown in Figure 2.2, the aaGAS approach produces volatility patterns that differ substantially from those of the standard and accelerating GAS models during the COVID-19 pandemic. While the standard GAS and aGAS specifications do not capture the shock induced by pandemic restrictions, the aaGAS model estimates notably higher volatility, consistent with the interpretation of a transitory shock. In the long run, the HLW constant estimate

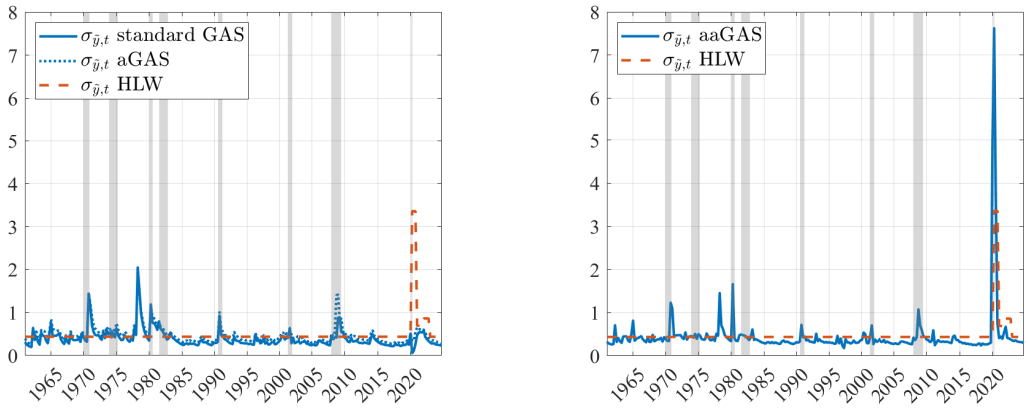

Figure 2.2: The solid blue lines show the output gap time-varying volatilities estimated by the standard GAS (left) and augmented-accelerating GAS update (right). The red dashed lines display the HLW estimates adjusted in the COVID-19 pandemic by the HLW (2023) methodology. Shaded vertical areas indicate US recessions as dated by the National Bureau of Economic Research (NBER).

interpolates between the score-driven volatilities, while during the COVID period it lies between the GAS and aaGAS estimates.

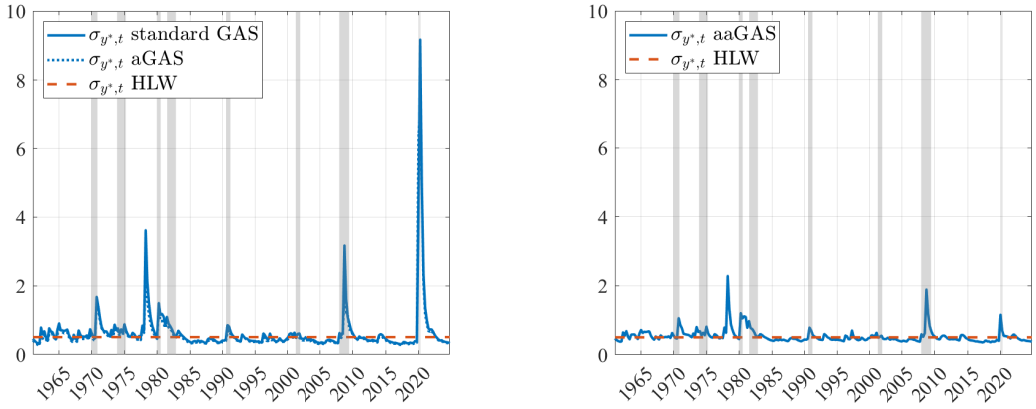

Figure 2.3: Potential output time-varying volatilities estimated by the standard GAS (left) and augmented-accelerating GAS update (right). The red dashed lines show the time-invariant counterparts estimated by HLW. Shaded vertical areas indicate US recessions as dated by the National Bureau of Economic Research (NBER).

Introducing fully state-dependent persistence in output gap volatility does not limit its impact to a single parameter. This is evident in Figure 2.3, which shows the difference between the conditional volatility estimates of the potential output obtained from the score-driven models. While the standard and accelerating GAS update allows pandemic-related innovations to affect the signal, the aaGAS updating mechanism mitigates the impact of these shocks on state volatility by down-weighting the associated scores through Eq. (2.12) and Eq. (2.13) This behavior is better understood when considering that pandemic-specific trajectories in volatility dynamics result in a different evolution of the signal-to-noise ratio, that is, the ratio of output gap to potential output volatility, a critical factor in estimating the state through the KF. As shown in the bottom panel on the left-hand side of Figure 2.4, when the model is estimated using the standard and accelerating GAS model, the ratio increases significantly due to the COVID-19 pandemic shocks, which considerably affect the path of potential output.

In contrast, as will be shown in Section 2.3.6, the aaGAS model discounts the weights of the same shocks in the KF through the signal-to-noise ratio (bottom panel on the right-hand side of Figure 2.4), thus minimizing the impact of transitory shocks on the state variable and leaving the potential output nearly unaffected. In addition, the bottom-right panel shows that, while the estimated trajectories obtained from the aaGAS and the HLW estimates follow a broadly similar pattern, their timing differs slightly when compared. It should be noted that the delayed response of the latter results from a COVID-related modeling choice by Holston et al. (2023), which allows the volatility of the underlying output gap to vary annually, but only starting in the second quarter of 2020, despite the fact that the US economy had already been affected by the COVID shock in the first quarter of 2020.

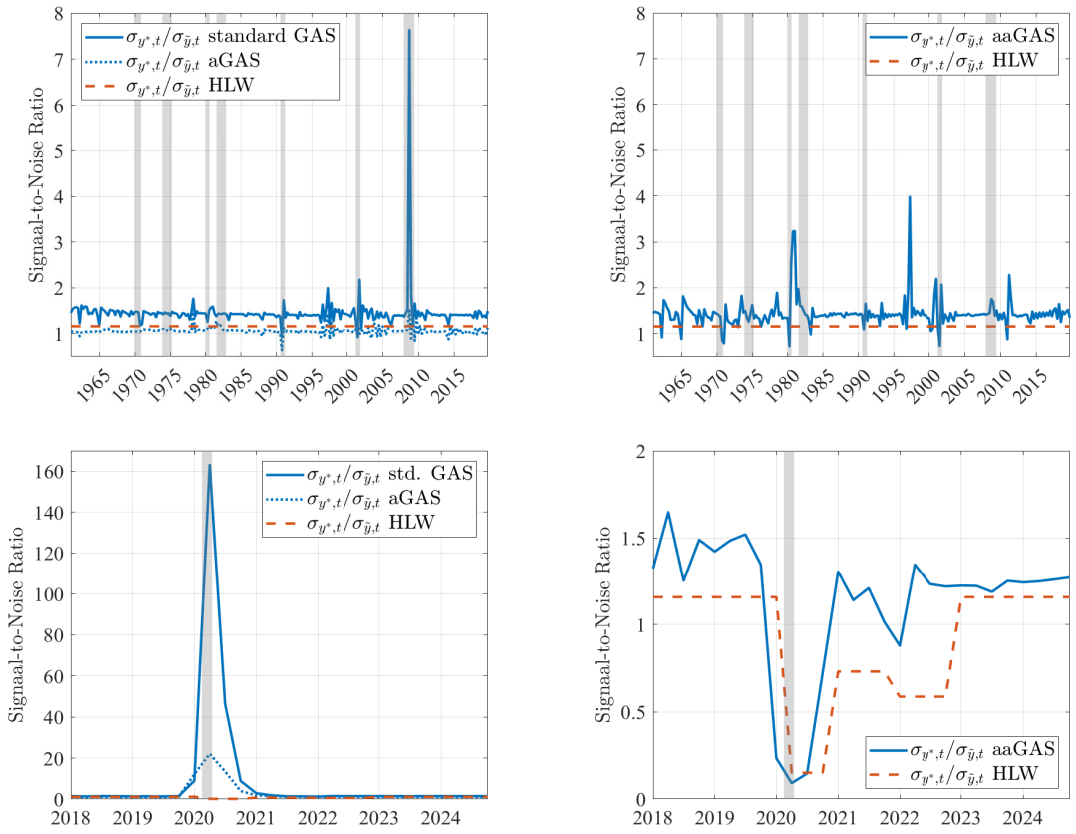

Figure 2.4: The solid blue lines show the signal-to-noise ratio (output gap volatility to potential output volatility) estimated by the standard GAS (left-hand side column) and augmented-accelerating GAS update (right-hand side column) in the pre-pandemic interval (upper panels) and from 2018 onward (bottom panels). The red dashed lines display the HLW estimates adjusted in the COVID-19 pandemic by the HLW (2023) methodology. Shaded vertical areas indicate US recessions as dated by the National Bureau of Economic Research (NBER).

Moving forward, Figure 2.5 depicts the volatility dynamics of inflation, which are consistent with major inflationary episodes.[5] The increases observed in the 1970s align with the oil shocks and energy crisis, which triggered a period of persistently high inflation in the US. It was not until the mid-1980s that inflation volatility subsided, following the Volcker monetary tightening. This period of low volatility, known as the Great Moderation, came to an end with the GFC, when the

[5]Since the estimates of this time-varying parameter largely coincide across the two models, only the aaGAS estimates are reported.

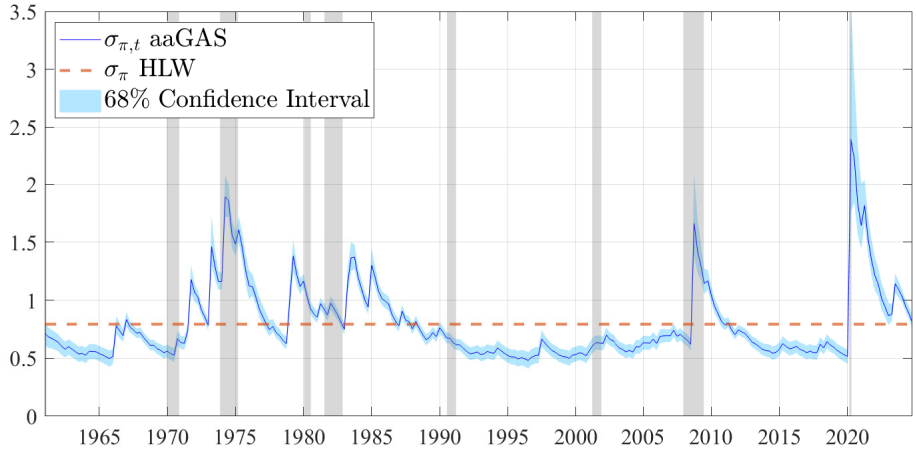

Figure 2.5: Inflation volatility estimated by the aaGAS model. The colored bands denote the 68% confidence interval corresponding both for filtering and parameter uncertainty. The red dashed lines show the time-invariant inflation volatility estimated by HLW (2017). Shaded vertical areas indicate US recessions as dated by the National Bureau of Economic Research (NBER).

standard deviation of inflation increased sharply. Finally, the inflationary shocks associated with the COVID-19 pandemic, followed by disruptions in energy markets in 2023, are also reflected in the evolution of the Phillips curve-related standard deviation.

We close the section by noting that the remainder of the analysis focuses primarily on the standard and aaGAS specifications for ease of exposition. This is because the accelerating GAS specification proves insufficient during the COVID interval to fully capture the volatility shift, as suggested by its signal-to-noise ratio (bottom-left panel of Figure 2.4), and constitutes an intermediate specification between the standard GAS and aaGAS models, as indicated by model fit.

2.3.5 Dynamic IS and Phillips curves

This section evaluates the dynamics of the parameters that drive the structural relationships of the IS and the Phillips curve. As the evolution of the two time-varying parameters in the two score-driven models considered takes a comparable path, only the results obtained by estimating the aaGAS model are reported.

Figure 2.6 displays the dynamics of the slope parameters of the IS and Phillips curves over the analyzed interval, respectively. Specifically, $a_{\tilde{y},t}$ corresponds to the output gap sensitivity to changes in the real interest rate gap, and $b_{\pi,t}$ determines the linkage between the output gap and the inflation rate. Although a causal interpretation of the endogenous relationships requires the identification of individual structural shocks, the main linkages can be summarized as follows. When the real funds rate is above (below) the natural rate of interest, the positive (negative) real rate gap reduces (increases) economic activity. The resulting lower (higher) output gap is, in turn, associated with deflationary (inflationary) conditions.

The evolution of the IS slope parameter in the left-hand panel confirms that the relationship between the real interest rate gap and the output gap underwent considerable structural changes over the analyzed interval. Specifically, until the economic recovery from the 1990s recession, the IS slope appears unstable and relatively steep. However, in line with the unfolding Great Moderation, characterized

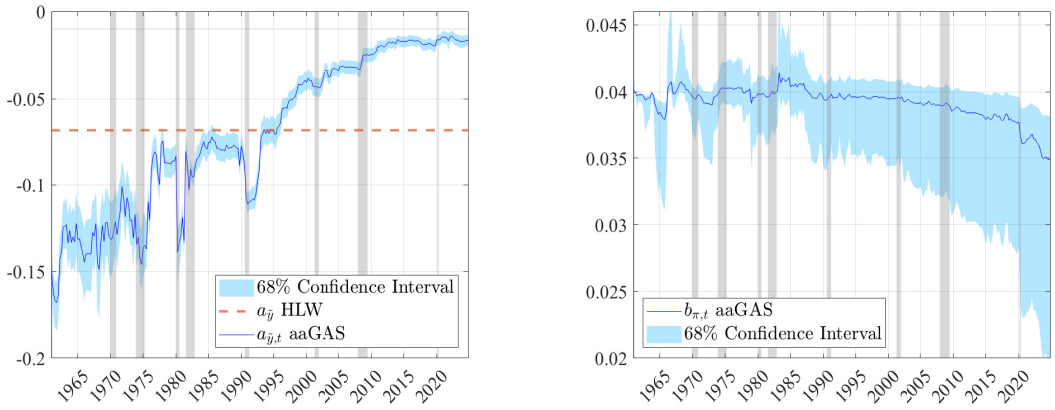

Figure 2.6: Dynamic slopes of the IS curve (left-hand side) and Phillips curve (right-hand side) based on simulation output. The red dashed line shows the time-invariant slope of the IS curve estimated by HLW (2017). The Phillips curve slope parameter estimated by HLW (2017) as 0.08 is not displayed to improve readability. The solid blue line corresponds to the 50th percentile of the simulated parameter. Shaded vertical areas indicate US. recessions as dated by the National Bureau of Economic Research (NBER).

by downward shifts in economic volatility and consolidating monetary policy, the slope gradually became flatter, reaching its current relatively stabilized level after the Great Recession.

Although the slope dynamics of the Phillips curve in the right-hand side panel follow a downward trend, the median of the simulated parameter path remains confined to a narrow range due to the low weight placed on its score innovation, $b_{\lambda,\pi}$, as reported in Table 2.4. This finding suggests that the relationship between the output gap and the inflation rate has been relatively stable in the analyzed interval, consistent with studies reporting little or no significant change in the slope of the Phillips curve in recent decades (e.g., Hazell et al., 2022). Nevertheless, the asymmetrical widening of the confidence bands over the last two decades makes it difficult to rule out the flattening Phillips curve hypothesis.

2.3.6 The r-star and the output gap

To open the discussion, we compare the dynamics of the output gap around the COVID pandemic obtained by the two GAS specifications discussed in Section 2.3.4, with those implied by the HLW estimates. As Figure 2.7 illustrates, the output gap filtered state trajectories under the two score-driven models diverge markedly from the onset of the COVID-19 pandemic. Interestingly, from this point onward, the median realization of the simulated state under the standard GAS model deviates from its filtered counterpart, resulting in more comparable simulated paths across the considered models. However, under the aaGAS model, consistent with the significance of the weighting parameters related to the standard deviation of the output gap reported in Table 2.3, the narrower confidence bands, with the closer alignment between the simulated and filtered states, suggest that introducing time-varying persistence in the volatility of the output gap substantially improves both the precision and the real-time reliability of the estimates.

While the aaGAS output gap exhibits the expected behavior as the COVID-19 pandemic evolves, the decline in its magnitude differs markedly from that estimated by HLW. Figure 2.8 presents the distinct estimates of the underlying states that, by definition, give rise to the observed dynamics. Notably, the natural rate of

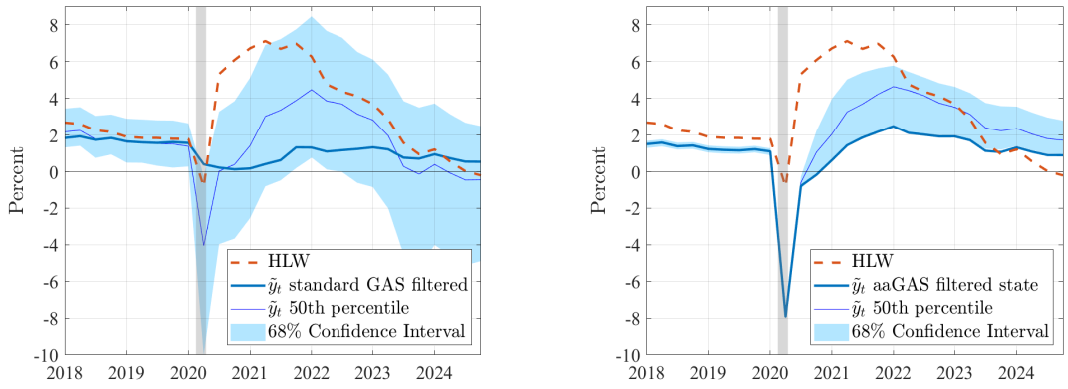

Figure 2.7: Estimations of the output gap around the COVID pandemic using the model with standard GAS (left-hand side) and augmented-accelerating GAS (aaGAS) update (right-hand side). The colored bands denote the 68% confidence interval accounting for both filtering and parameter uncertainty. The thin solid blue line indicates the 50th percentile of the simulated states, while the thick blue line shows the real-time filtered estimates. Shaded vertical areas indicate US recessions as dated by the National Bureau of Economic Research (NBER).

output moderately declines in all models considered following the initial impact of the pandemic, but diverges noticeably thereafter. Although the y-star estimated by the standard GAS and HLW models falls more sharply in 2020Q2, the former rebounds significantly, whereas the latter continues along a shifted trend as the economic impact of the pandemic unfolds. In contrast, the aaGAS model interprets pandemic-related shocks as transitory, resulting in a potential output estimate that follows a moderately shifted trend with a milder slope. Given the primary objective of the Kalman filter in this study, the substantial shift in the standard GAS and HLW state estimate appears to be difficult to reconcile with the typically persistent nature of potential output. In contrast, the aaGAS scenario, which shows a trend with a moderate downward shift, with increased uncertainty around the output gap, aligns more closely with the somewhat neutralizing forces on productivity experienced after the initial shock of the COVID-19 pandemic.[6]

Looking at the output gap over the full interval, Figure 2.9 reveals that the declines in output gap at the aaGAS model are generally deeper than the HLW estimations. The fall in the output gap is especially pronounced following the Volcker shock, which proceeds with a slower rebound than the HLW estimation. This mismatch arises from the higher potential output the score-driven model gives, likely reflecting its muted sensitivity to transitory shocks consistent with the structural assumption. These characteristics suggest that models with time-varying volatilities improve the effectiveness of filtering out temporary disturbances when estimating state variables. In addition, the resulting asymmetry, which, except for the high-inflation trend episodes in the 1960s and 1970s and the post-COVID period, dominates the analyzed interval, is somewhat compatible with the "plucking model" by Friedman (1993), where the economy fluctuates under its full potential ceiling.[7]

[6]For example, Bloom et al. (2025) find little lasting impact on aggregate total factor productivity with significant heterogeneity across firms and sectors in the US. Similarly, Fernald and Li (2022) report a relatively modest decline in potential output in the post-COVID interval, driven by offsetting effects across different channels.

[7]The plucking framework has been studied recently both theoretically (Suah, 2024) and empirically (Dupraz et al., 2019).

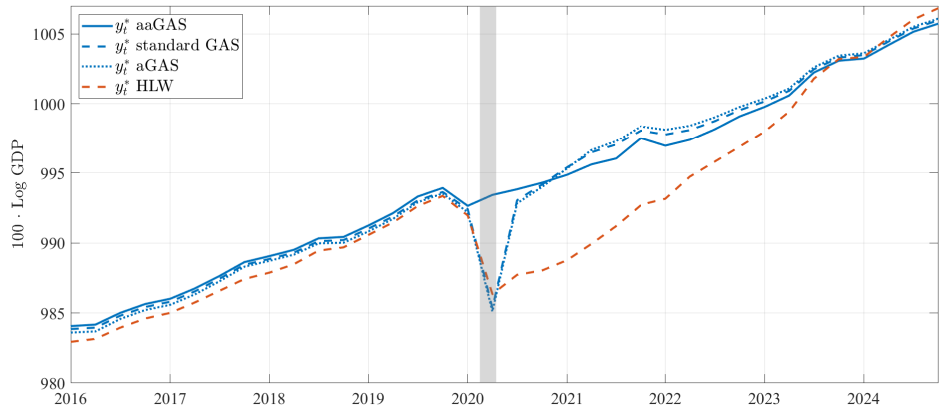

Figure 2.8: Estimation of the potential output by different models around the COVID-19 pandemic.

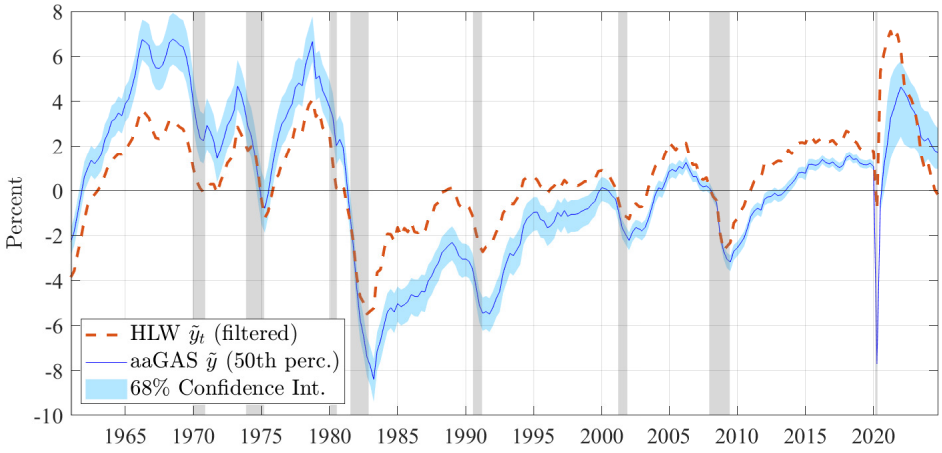

Figure 2.9: Output gap 50th percentile simulation projection by the aaGAS model with the corresponding confidence band and the benchmark HLW filtered estimate. Shaded vertical areas indicate US recessions as dated by the National Bureau of Economic Research (NBER).

Turning to the central interest of the study, Figure 2.10 shows the r-star estimated by the aaGAS specification with a confidence interval accounting for both parameter and filtering uncertainty.[8] As the plot displays, the natural rate of interest follows a downward trend through the analyzed interval and, as shown in Figure 2.11, its real-time estimate is characterized by a relatively lower variability than the corresponding HLW estimate. Most importantly, the decline during the GFC is less pronounced in our estimation, and even afterwards, it tends to remain in a higher range than the HLW estimate. The distinct trajectory of the aaGAS r-star, as shown by the two-sided estimates in Figure 2.11, becomes even more evident when comparing its smoothed values given by the KF recursion using the full information set. However, considering the ongoing debate about the current location of the r-star, the most intriguing finding is perhaps the observed divergence in the trends of the two real-time estimates.

As Figure 2.12 shows, the primary source of the estimation discrepancy is the fundamentally different paths of the *other factors* component, z_t. Although the

[8]Except for the potential output, we report only the state estimations of aaGAS as these results are robust over the two specifications.

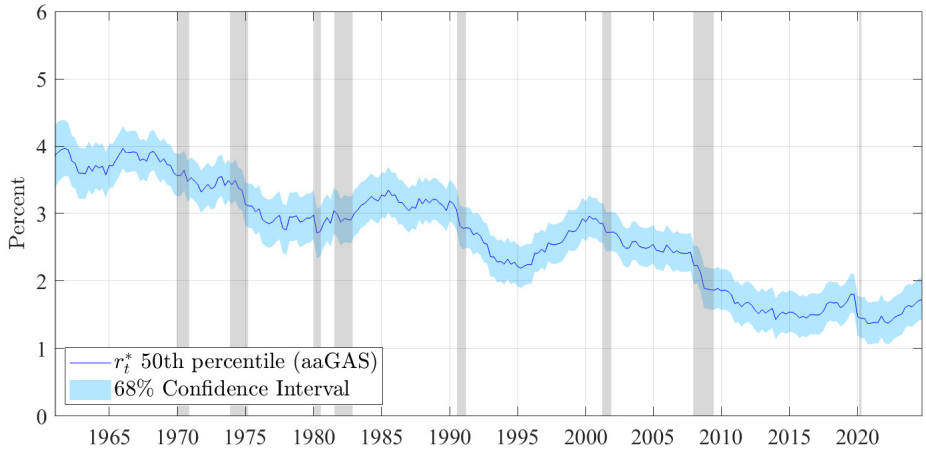

Figure 2.10: Simulation results for the natural rate of interest using the augmented-accelerating GAS (aaGAS) model. The colored bands denote the 68% confidence interval corresponding both for filtering and parameter uncertainty. Shaded vertical areas indicate US recessions as dated by the National Bureau of Economic Research (NBER).

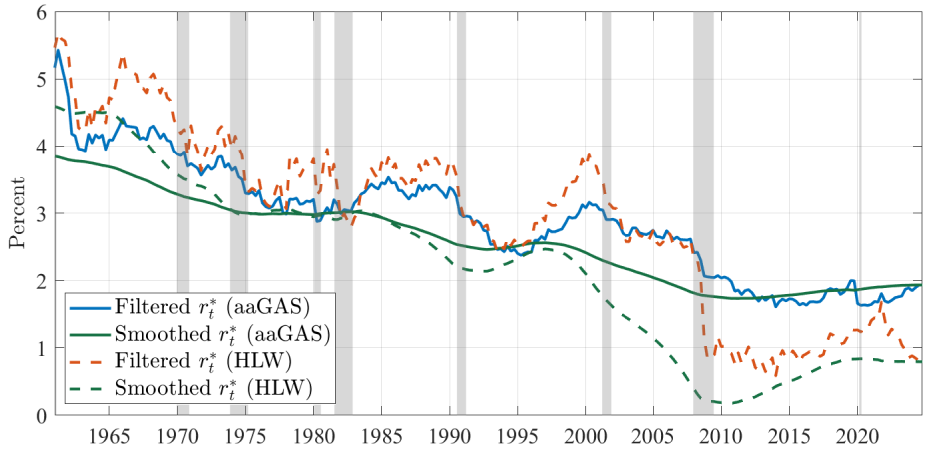

Figure 2.11: The solid blue line represents the filtered (one-sided) estimates of r-star using the aaGAS model, while the red dashed line shows the corresponding estimates from the HLW approach. The solid green line depicts the smoothed (two-sided) estimates of the natural rate of interest under the aaGAS specification, and the green dashed line shows the smoothed estimates based on the HLW model. Shaded areas indicate US recessions as dated by the National Bureau of Economic Research (NBER).

descent in z_t estimated by HLW accelerates in the post-pandemic interval, it shows a limited variation in our output. Notably, the two estimates depart significantly after the GFC, with a widening divergence over the rest of the interval. In contrast, as Figure 2.13 shows, the other determinant of the r-star, the trend growth rate, is characterized by substantially narrower confidence bands and, except for the post-GFC interval, a closer match to the HLW estimate. Overall, these observations, in line with the associated standard errors reported in Table 2.3, make it evident that most of the uncertainty surrounding the r-star and its higher trajectory over the past two decades relative to the HLW results stems from the estimation of the z-factor. At the same time, as discussed previously, the notably lower standard errors confirm that our models provide a more precise estimate of z_t compared to HLW, which largely accounts for the reduced uncertainty of our natural rate of

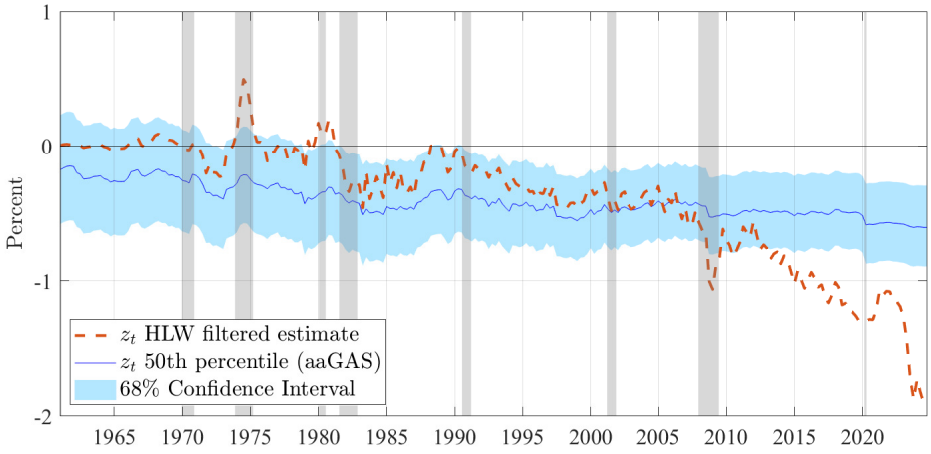

Figure 2.12: Estimation of the "other factors" component (z_t) by the model with augmented-accelerating GAS (aaGAS) update and with HLW 2023 configuration. The colored bands denote the 68% confidence interval corresponding both for filtering and parameter uncertainty. Vertical shadows indicate recessions as identified by the National Bureau of Economic Research (NBER).

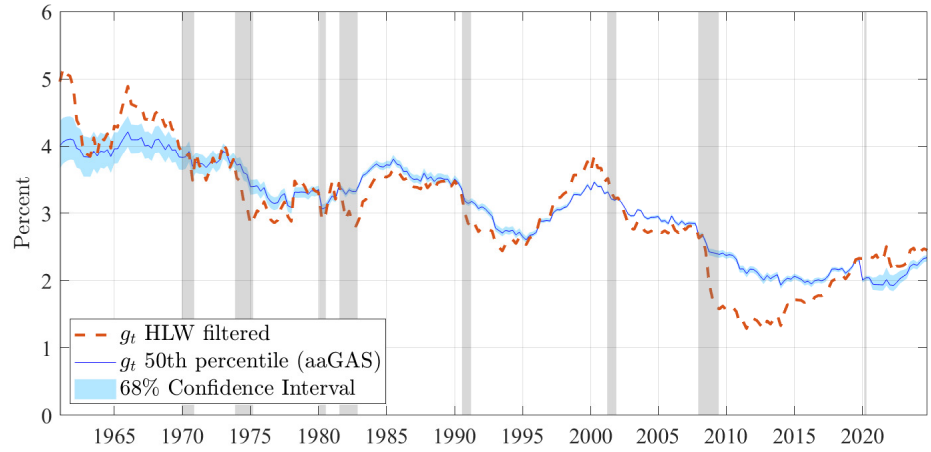

Figure 2.13: Estimation of the trend growth rate by the model with augmented-accelerating GAS (aaGAS) update and by HLW (2017). The colored bands denote the 68% confidence interval corresponding both for filtering and parameter uncertainty. Vertical shadows indicate recessions as identified by the National Bureau of Economic Research (NBER).

interest estimates.

As discussed above, the key component in the KF algorithm is the Kalman gain matrix, which assigns weights to the prediction errors and determines the latent state. Therefore, it is instructive to examine the evolution of its elements to understand how the score-driven framework influences the estimation via this instrument. We collect the main elements of the Kalman gain matrix in Figure 2.14, which shows how the weight of the information content in the IS curve (left-hand side column) and the equation of the Phillips curve (right-hand side column) are allocated through the Kalman filter recursion. While the weights associated with the potential output (upper two panels) and thus determining the output gap oscillate around the weight obtained from the HLW model, the rest of the panels show a different picture. Notably, the bottom panels of Figure 2.14 show that the Kalman gains related to the z-factor differ considerably between the two models. Specifically, apart from the periods affected by the COVID-related adjustments,

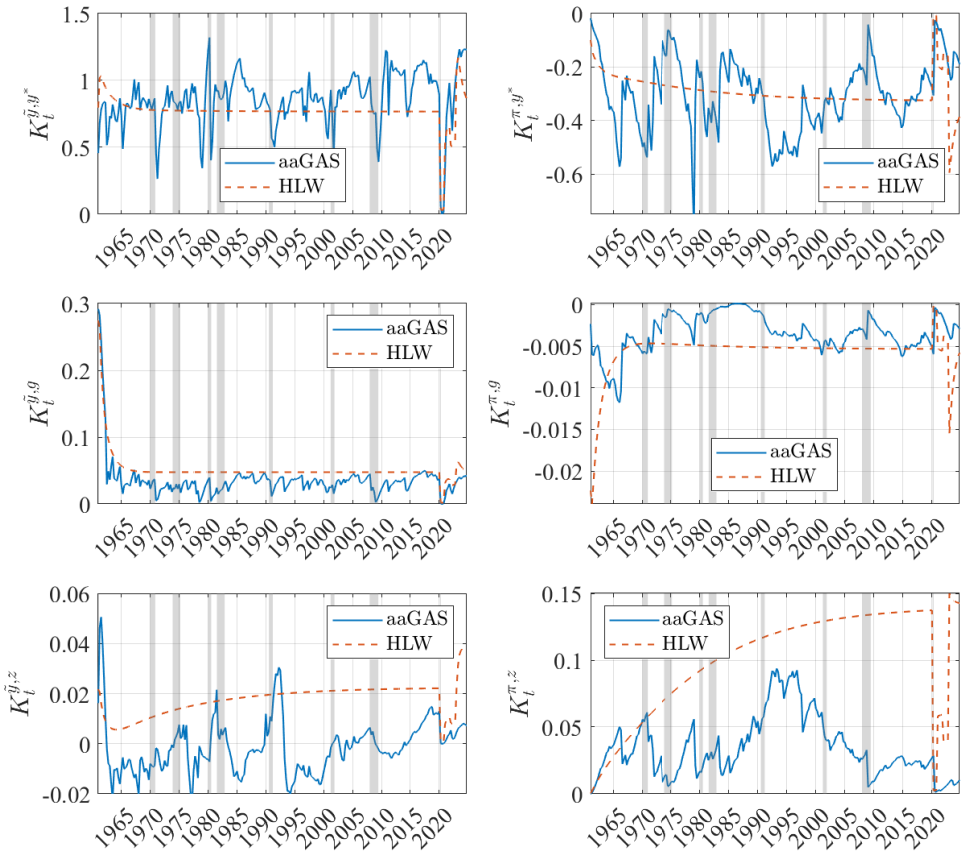

Figure 2.14: Kalman gains (K_t) associated with different structural equations and state variables. The panels display, from left to right and top to bottom: (1) Kalman gain associated with the potential output and IS equation, (2) Kalman gain associated with the potential output and Phillips curve, (3) Kalman gain associated with the trend growth rate and IS equation, (4) Kalman gain associated with the trend growth rate and inflation dynamics, (5) Kalman gain for the z-factor and IS relation, and (6) Kalman gain for the z-factor and Phillips curve equation.

the Kalman gains obtained from the HLW model are consistently higher than those from the aaGAS specification. Altogether, the high flexibility provided by the score-driven setup, as opposed to the relatively rigid and larger weights obtained from the static parameter model for the two components of the r-star, g_t and z_t, led to our higher and generally more stable natural rate of interest estimates over the past two decades, materializing primarily through the estimation of the z-factors.

2.4 Alternative identification and extensions in the Laubach–Williams framework

Our analysis shows that the r-star estimates depend critically on the "other factors" component, z_t. Its non-trivial contribution to the trajectory of r-star, together with the substantial ambiguity surrounding its interpretation and estimation, motivates alternative identification strategies and extensions of the baseline LW framework. While full identification is unlikely given the inherently composite nature of this factor, increasing the share of its economically interpretable component offers a potential way to improve both identification and the precision of the estimated

natural rate.

Given the relevance of uncertainty and risk-taking behavior for real economic activity and their implicit role in the present analysis, integrating financial risk-related dimensions constitutes a natural refinement of the LW framework and its identification strategy. This extension is consistent with a long-standing line of research on the role of financial factors in business fluctuations. Such approaches either remained outside the mainstream (e.g., Minsky, 2015; Kindleberger, 2000) or predate modern formal macroeconomic frameworks (e.g., Fisher, 1932). However, they have gradually been incorporated into more standard models, prominently through neoclassical frameworks augmented with financial frictions (e.g., Bernanke and Gertler, 1989; Bernanke et al., 1999).

The Global Financial Crisis and its aftermath renewed interest in incorporating financial frictions into New Keynesian models (e.g., Curdia and Woodford, 2010, 2011, 2016; Gertler and Kiyotaki, 2010; Gertler and Karadi, 2011). These models provide a natural theoretical foundation to link financial factors to the "other factors" component. In this class of models, spreads, external finance premia, and intermediary risk premia enter the IS equation as time-varying wedges, reflecting deviations between the policy rate and the effective rate relevant for consumption and investment decisions. Given that the natural rate of interest is defined as the real rate consistent with a zero output gap, persistent movements in these wedges shift the equilibrium intertemporal condition and, in turn, can affect the level of the natural rate of interest. Consequently, persistent changes in financial conditions driven by underlying risk attitudes can be interpreted as a determinant of the natural rate of interest, providing a more economically grounded decomposition.

A natural candidate for an external proxy for this purpose is the risk premium, defined as the compensation investors require for bearing risk above the risk-free rate. Yet, aggregate risk premium measures often conflate multiple components, obscuring underlying heterogeneity across the financial sector. Therefore, to mitigate the impact of idiosyncratic variation and avoid reliance on a single aggregate indicator, one can construct a *risk premium indicator* as the *first principal component* based on bank-level measures. Specifically, this approach would exploit cross-sectional information in intermediary balance sheets to extract the common variation in financial conditions. Consequently, it is better aligned with the theoretical drivers of time-varying financial wedges and provides a more informative proxy for the underlying financial conditions relevant for aggregate demand.

This interpretation is consistent with Bauer et al. (2023), where, based on standard asset pricing theory, risk appetite is defined as the common movement in risk premia and risky asset prices across multiple markets. In this sense, the comovement of risk premia across entities within a given sector can be interpreted as reflecting changes in the overall risk appetite for that sector. This approach is further supported by the granular perspective outlined in Gabaix (2011), which shows that when the distribution of firm sizes is fat-tailed, idiosyncratic shocks at the level of large firms can drive aggregate fluctuations. Since a few large banks own a large share of the aggregate assets in the US banking sector, this implies that common variation in balance sheet conditions across major institutions can capture economically meaningful aggregate dynamics. Given this, extracting a principal component from balance sheet indicators of banks above an economically relevant size threshold provides a tractable way to capture the common variation across intermediaries. The resulting measure serves as a sector-wide proxy for risk-taking that is both empirically grounded and consistent with the presence of

granular effects.

Building on the above, the extracted factor p_t can be interpreted as a summary indicator of sector-wide financial conditions associated primarily with intermediary balance sheet conditions, while also reflecting broader private sector financial conditions. However, for the purpose of linking financial factors to the natural rate of interest, it is useful to distinguish between its persistent and cyclical components. Therefore, the financial factor is decomposed into a component p_t^*, capturing persistent changes in underlying financial conditions, including structural and behavioral factors such as shifts in risk attitudes, private-sector deleveraging (Eggertsson and Krugman, 2012), long-term regulatory changes, and financial innovation, and a gap $\tilde{p}_t$, reflecting cyclical fluctuations in financial conditions.

This decomposition is related to the distinction between the financial cycle and the financial gap (e.g., Borio, 2014), in that it separates short-term fluctuations from more persistent financial dynamics. The cyclical component $\tilde{p}_t$ captures short-term deviations in the effective cost of capital, consistent with time-varying financial wedges in the IS equation. In contrast, the persistent component p_t^* contributes to shifts in the natural rate of interest.

Specifically, the definition of latent states is modified and extended by the new measurement p_t:

$$p_t = \tilde{p}_t + p_t^*, \tag{2.27}$$

$$\tilde{y}_t = y_t - y_t^*, \tag{2.28}$$

$$r_t^* = g_t + z_t + p_t^*, \tag{2.29}$$

where p_t is the *standardized first principal component of risk premia across financial intermediaries*. The unit-level risk premium is defined as the spread between the lending rate and the effective federal funds rate, capturing bank-specific financial conditions reflected in credit risk, funding costs, intermediary margins, and the cost of external finance faced by borrowers. The lending rate is determined by dividing interest income by gross loans retrieved from commercial and investment banks' quarterly balance sheets and income statement data. The natural rate of interest is augmented to include the persistent component p_t^*, while the cyclical component $\tilde{p}_t$ enters the IS relationship in Eq. (2.2) as a financial wedge scaled by the parameter γ:

$$\tilde{y}_t = a_1\tilde{y}_{t-1} + a_2\tilde{y}_{t-2} + \frac{a_{\tilde{y},t}}{2}\sum_{j=1}^{2}\left[(r_{t-j} - r_{t-j}^*) + \gamma\tilde{p}_{t-j}\right] + \varepsilon_{\tilde{y},t}. \tag{2.30}$$

With the transition equation for p_t^* complementing the system of Eqs. (2.5)–(2.7):

$$p_t^* = p_{t-1}^* + \varepsilon_{p^*,t} \qquad\qquad \varepsilon_{p^*,t} \sim \mathcal{N}(0, \sigma_{p^*}) \tag{2.31}$$

Overall, this extension offers a structured way to incorporate financial factors into the LW framework and to provide a more interpretable decomposition of the natural rate of interest. However, its implementation remains subject to measurement and identification challenges.

An alternative identification strategy for the Phillips curve relationship could also be considered. Specifically, relaxing the assumption of strictly backward-looking price-setting behavior would align the model more closely with the New Keynesian framework. For this purpose, the extended series of Clark and Doh

(2014) on long-term inflation expectations, employed in Chapter 3, could replace the term $(1 - b_1)\pi_{t-2,4}$ in Eq. (2.3), yielding:

$$\pi_t = b_1\pi_{t-1} + (1 - b_1)\pi^E_{t-1} + b_{\pi,t}\tilde{y}_{t-1} + \varepsilon_{\pi,t}, \tag{2.32}$$

where π^E_{t-1} denotes lagged long-term inflation expectations.

In summary, these extensions suggest a broader strategy of sharpening the identification of the LW framework by enriching the information set used, including micro-level data. Further research should assess the joint performance of these extensions, their robustness across specifications, and their implications for the precision and interpretability of the estimated natural rate of interest.

2.5 Conclusions

Determining the natural rate of interest has been an inspiring subject of academic discourse since the concept was first introduced over a century ago. This chapter contributes to this line of research by making the parameters of the Laubach-Williams semi-structural model time-varying, with their dynamics being governed by the conditional likelihood score. In addition, motivated by the econometric challenges initiated by the pandemic, we introduced an augmented accelerating version of the GAS (aaGAS) updating mechanism.

The output gap estimated by the aaGAS suggests that the modified model effectively handles large volatility shifts due to extreme events. Furthermore, the higher estimate of potential output relative to HLW suggests that the proposed models are more efficient at disentangling transitory and highly persistent shocks. Regarding the time-varying structural relationships, the IS slope was highly unstable until the beginning of the Great Moderation, gradually flattening during this period and stabilizing at a low level after the Great Recession. As for the Phillips curve, while we find little evidence of time variation in its slope parameter, the pronounced downside uncertainty in the simulated parameter limits any decisive conclusions regarding the stability of the slope over the past decades. Most importantly, the r-star estimates exhibit substantially higher precision and reduced variability, along with a flatter downward trend over the analyzed interval compared to the benchmark HLW model. In addition, we find that the primary source of the recent divergence in the trajectories and the uncertainty surrounding the natural rate of interest lies in the estimation of the "other factors", z. These state estimates not only deviate substantially from those of HLW but also account for most of the improvement in r-star estimation precision, explaining the higher natural rate of interest—ranging between 1.5% and 2%—observed over the past two decades. Furthermore, the current upward trend in both the filtered and smoothed estimates, in contrast to HLW model output, indicates that the natural rate of interest has already passed its historically low level.

Given the high uncertainty and the ambiguous definition of the z-factor, future research could focus on improving its refined identification and developing more precise specifications. Moreover, considering the low-frequency nature of current estimations and the delayed availability of relevant data, a model that utilizes higher-frequency data and proxy variables could offer a more timely and accurate assessment.

In summary, the lower variability of the r-star estimate, along with the reduced standard errors of the underlying state estimates, implies that the proposed models are more adaptable to sudden changes in the economic environment and, thus,

more effective at identifying transitory shocks than the standard LW methodology. Based on a purely statistical, data-driven approach, the GAS (score-driven) models, therefore, provide a valid empirical alternative for estimating the natural rate of interest in real-time, demonstrating a clear advantage in estimation precision. In addition, the aaGAS specification further improves the precision of the r-star estimate, alleviating a long-standing limitation of models based on the LW methodology. Moreover, in light of the occurrence of extreme events in both macroeconomic and financial contexts, the aaGAS offers a viable extension of the standard GAS framework to more adequately model volatility persistence.

Chapter 3

Inflation Dynamics and the Phillips Curve in the US: The Smoothed Dynamic Multiple Quantile Model

This chapter is based on "Inflation Dynamics Beyond the Mean: A Time-Varying Multiple Quantile Phillips Curve with Quantile-Heterogeneous Slopes" by Tibor Pál (working paper, 2026).

Abstract

This chapter studies the dynamics of the conditional distribution of US inflation within a New Keynesian Phillips curve framework. Conditional inflation quantiles are jointly modeled using a dynamic multiple quantile specification. To accommodate time variation and quantile heterogeneity, the smoothed Dynamic Multiple Quantile model is introduced that allows the Phillips curve slope to vary over time and across the distribution. We document pronounced nonlinearities in the Phillips curve and in the relative importance of backward and forward-looking inflation expectations. The estimated upper-tail slope is consistently flat across specifications, indicating pronounced asymmetry across quantiles. The entire cross-quantile profile evolves over time, implying a time-varying Phillips curve shape. Periodic flattening of the Phillips curve, particularly in the lower tail, is consistent with the missing disinflation following the Great Recession and in the post-pandemic period. Lastly, the filtered conditional inflation quantiles, together with higher-order moments, provide a quantitative measure of inflation tail risk.

3.1 Introduction

Price stability is a central objective of monetary policy, and the evaluation of inflation risk is an important input to fulfill this mandate. Understanding the conditional distribution of inflation, its determinants, and their dynamics is therefore crucial for policy decisions.

Since the seminal contribution of Phillips (1958), growing attention has been devoted to identifying the key drivers of inflation dynamics, particularly the relative roles of forward- and backward-looking price-setting behavior and the strength of the relationship between real activity and inflation. Despite substantial methodological progress, empirical findings remain mixed, and a consensus has yet to emerge. This is particularly concerning given the widespread imposition of this theoretical link in studies aiming to identify quantities related to equilibrium or long-term concepts such as the output gap or the natural rate of unemployment.

Given this context, this chapter contributes to the literature on inflation dynamics by jointly modeling multiple quantiles of the US inflation within a framework motivated by the New Keynesian Phillips Curve. Building on the Dynamic Multiple Quantile approach, the model allows the effects of inflation determinants to vary both over time and across quantiles corresponding to predefined probability levels.

Traditionally, financial econometricians have focused on studying the behavior of conditional distributions in the tails, a focus that has become increasingly relevant in macroeconomic contexts. Building on the well-established concept of value at risk (VaR), a risk metric that quantifies the highest potential loss of a financial portfolio or asset within a specific time frame at a chosen probability threshold, research into the tail-end behavior of macroeconomic variables has attracted considerable interest in recent years.

At-risk frameworks, such as growth-at-risk (Adrian et al., 2019), inflation-at-risk (López-Salido and Loria, 2024), and debt-at-risk (Furceri et al., 2025), provide valuable tools to assess the conditional distribution of macroeconomic variables. However, these studies typically rely on time-invariant reduced form quantile regressions, in which conditional distributions are linked to a set of explanatory variables. Specifically, they are based on static specifications with parameters that are constant within each estimation period and exhibit limited state dependence.

Consequently, while these models provide insight into how current conditions map into future distributions, they remain limited in their ability to capture the time-varying and state-dependent nature of quantile dynamics. These limitations are especially pronounced when modeling inflation dynamics, where persistence is a well-documented feature implying a gradually evolving conditional inflation distribution through the propagation of past shocks.

Therefore, in contrast to the standard quantile regression approach, and aligned with a time-series setting, we consider a dynamic quantile model in which the quantile of the distribution of the random variable under study is conditioned on the available prior-period information, $p(y_t \mid \mathcal{I}_{t-1})$, where $\mathcal{I}_{t-1}$ contains the entire history of y_t up to time $t-1$, thereby introducing an explicit backward-looking information component.

The proposed model builds on the Dynamic Multiple Quantile (DMQ) framework of Catania and Luati (2023), with two key departures. First, in line with the New Keynesian Phillips Curve (NKPC), additional variables are included in the specification, namely proxies for marginal cost and inflation expectations. Second,

reflecting theoretical considerations underlying the dynamic Phillips curve relationship, the slope parameter varies over time and across quantiles. In addition, inflation inertia and the weight on inflation expectations vary over time, making explicit the conditional role of backward- and forward-looking components.

An inherent feature of the DMQ model is its reliance on a large set of quantiles to provide sufficient information for updating the dynamic quantiles. Introducing quantile-specific parameters in the original framework therefore results in a substantial increase in model dimensionality. This introduces a trade-off between retaining sufficient information to update the dynamic quantiles and maintaining parsimony in the parameter space. In the proposed framework, this trade-off is relaxed by specifying parameters at a selected subset of probability levels and interpolating them at the remaining quantiles. Therefore, the developed DMQ specification allows the parameters associated with inflation determinants to vary over time and across quantiles while keeping the overall parametrization tractable. The model is interpreted as an observation-driven approximation and is labeled as the Smoothed DMQ (sDMQ) model, where *smoothed* refers to the use of monotone interpolation to construct a continuously differentiable parameter curve across probability levels, thereby ensuring smooth parameter variation over the conditional inflation distribution.

The main contributions of the study are as follows. First, we generalize the DMQ framework by allowing parameters to vary jointly across quantiles and over time, yielding a time-varying and heterogeneous DMQ specification. Second, we employ this framework to analyze US inflation dynamics using specifications motivated by the New Keynesian Phillips Curve. Third, we provide empirical evidence of state-dependent inflation inertia, the role of inflation expectations, and nonlinear Phillips curve dynamics across both time and the conditional distribution, including a time-varying cross-quantile shape and a persistently flat slope in the upper tail. Fourth, we demonstrate that the choice of expectation proxy is crucial for identification: short-run expectations generate richer higher-order dynamics, whereas more anchored long-run expectations attenuate tail behavior.

A key challenge in estimating the New Keynesian Phillips curve using limited-information econometric methods is the limited identifying power of economic aggregates, as emphasized by Mavroeidis et al. (2014). This chapter shows that incorporating distributional information and time variation can provide additional insights into inflation dynamics beyond conventional aggregate approaches.

Compared to a traditional Phillips curve model, the proposed framework offers several advantages. First, it captures time-varying and heterogeneous responses of inflation to its determinants. For example, the Phillips curve relationship at upper quantiles can differ markedly from that at lower quantiles, which is relevant for monetary policy. Second, the model characterizes the dynamics of the conditional inflation distribution, providing information beyond the first two moments that is useful for monetary policy and risk management.

The empirical analysis, based on quarterly US core inflation, reveals clear heterogeneity in Phillips curve relationships across quantiles. The results indicate an asymmetric Phillips curve that flattens in the upper tail and exhibits time variation concentrated in the lower and middle parts of the distribution. This profile evolves over time, with both its shape and magnitude shifting across macroeconomic regimes. Using alternative expectation measures, the short-run specification yields less persistent slope estimates, while the long-run specification delivers slopes with greater and more persistent variability. The overall slope profile and its time

variation remain robust across specifications.

During the 1970s, an increase in the weight on inflation expectations is consistent with weakly anchored expectations and the high-inflation regime of the Great Inflation. Following the Global Financial Crisis, the flattening of the Phillips curve in the lower tail, combined with strong inflation persistence, is consistent with the missing disinflation observed during the recovery. In the post-COVID period, further changes in the slope profile and a renewed role for expectations highlight the importance of state dependence and time variation in inflation dynamics.

The chapter is structured as follows. Section 3.2 presents the core econometric framework. Section 3.3 then introduces the benchmark empirical analysis, which closely follows the backward-looking NKPC framework considered in Chapter 2 in order to facilitate comparability across methodologies. Building on this benchmark, Section 3.4 progressively refines the specification to address its main limitations and to improve the representation of inflation dynamics within the sDMQ framework. Specifically, it incorporates forward-looking inflation expectations, thereby allowing for a hybrid NKPC specification with a time-varying relative weight of forward- and backward-looking inflation expectations. In addition, it integrates monotone piecewise cubic interpolation, additional identifying restrictions, and alternative measures of real marginal cost. The section also presents the main empirical findings, including the estimated Phillips curve slope dynamics, the relative contribution of backward- and forward-looking inflation expectations, and the evolution of the conditional inflation distribution and its implied moments. Finally, Section 3.5 concludes the chapter.

3.2 The econometric framework

The econometric framework for modeling inflation quantiles builds on the Dynamic Multiple Quantile (DMQ) model of Catania and Luati (2023). A key limitation of the DMQ model is its reliance on a large set of quantiles to update the time-varying quantile processes. When the focus is on studying dynamic relationships with semi-structural heterogeneity across the distribution, increasing the size of the jointly modeled quantile set reduces the flexibility of the model specification.

To address this shortcoming, the proposed model imposes a more parsimonious parametrization. Specifically, parameters of the explanatory variables are specified only at selected quantiles corresponding to predetermined probability levels and are subsequently interpolated for the remaining quantiles. Focusing on a reduced set of quantile processes facilitates the analysis of relationships across the distribution and opens the way to introduce time-varying parameters.

The section begins with a brief overview of the dynamic quantile Phillips curve model before outlining its main building blocks in the following subsections. The baseline DMQ model, its properties, and its limitations are then discussed in detail, motivating the subsequent modifications that lead to the proposed sDMQ model.

Lastly, the sDMQ model with time-varying parameters is presented, providing an effective framework for analyzing the dynamics of inflation quantiles. The methodology incorporates additional variables in line with a New Keynesian Phillips curve relationship, with the slope parameter varying across quantiles and over time.

3.2.1 The dynamic multiple quantile Phillips curve model

This section introduces the general structure for modeling the dynamics of the inflation distribution within a DMQ framework. The model is motivated by a New Keynesian Phillips Curve (NKPC), relating inflation to its expectations and allowing for a trade-off between the output gap and inflation.

The main departure from the baseline DMQ specification is the inclusion of exogenous variables whose effects on the modeled quantiles are allowed to be heterogeneous and time varying. This objective introduces a key difficulty. The original DMQ model requires joint modeling of a large set of quantiles to obtain sufficient information for the score-driven updates, thereby limiting the feasible parameter dimensionality.

The proposed approach addresses this econometric challenge by focusing on a subset of quantiles at selected locations of the conditional distribution. Throughout the remainder of the chapter, these quantiles of interest are referred to as *target quantiles*, and the parameters governing the structural relationships at the associated probability levels are referred to as *target parameters*. The parameter values at the remaining quantiles are obtained by interpolation, implying that these *auxiliary quantiles* are governed by *auxiliary parameters* that are deterministic, smooth functions of the target parameters. This construction restores a large set of quantiles, providing sufficient information for updating the target quantiles. As a result, the sDMQ specification enables the analysis of heterogeneous and dynamic responses of inflation to its determinants across different parts of the inflation distribution.

Quantiles divide the conditional distribution into regions such that a given proportion of observations falls below each quantile value at a specific probability level. In other words, the p-th conditional target inflation quantile $\hat{q}^{\tau_p}(\pi)$ is the inflation rate at which the conditional probability of inflation being less than $\hat{q}^{\tau_p}(\pi)$ equals τ_p. Meanwhile, the probability of inflation being greater than $\hat{q}^{\tau_p}(\pi)$ is equal to $1 - \tau_p$.

Consequently, a general NKPC-based specification for the conditional quantile function of inflation, $\hat{q}_t^{\tau_p}(\pi)$ is given by

$$\hat{q}_t^{\tau_p}(\pi) = \tilde{q}_t^{\tau_p}(\pi) + \kappa_{p,t}\, x_t + \delta\, \pi_t^E, \tag{3.1}$$

where $\hat{q}_t{}^{\tau_p}(\pi)$ denotes the pth conditional target inflation quantile at the probability level τ_p, given the information set available at time $t-1$. The term x_t represents a proxy for real marginal cost, with a time-varying slope parameter $\kappa_{p,t}$ specific to each quantile process. The variable π_t^E denotes inflation expectations with its weight parameter δ. The term $\tilde{q}_t^{\tau_p}(\pi)$ denotes the score-driven *dynamic quantiles*, each with its own updating process, described in the following section.

To conclude the section and for ease of reference, Table 3.1 summarizes definitions of the key quantile objects and parameters employed in the sDMQ framework.

Table 3.1: Notation and Definitions

Term	Notation	Definition
Target quantiles	$\hat{q}_t^{\tau_p}$	Directly modeled semi-structural conditional quantiles.
Auxiliary quantiles	$\bar{q}_t^{\tau_j}$	Quantiles not directly parametrized.
Dynamic quantiles	$\tilde{q}_t^{\tau_j}$	Score-driven quantiles driving the target dynamics.
Target parameters	$\hat{\kappa}_{p,t}$	Time-varying slope parameters at target quantiles.
Auxiliary parameters	$\bar{\kappa}_{j,t}$	Time-varying slope parameters at auxiliary quantiles.

3.2.2 The score-driven update of the DMQ model

Several features make the DMQ framework attractive for jointly studying the dynamics of the inflation quantiles. First, the model satisfies the non-crossing quantile criterion by construction. Second, it is consistent with the defining property of fixed quantiles, ensuring that in the limit the proportion of observations falling below the unconditional τ-level quantile converges to τ. Third, the model can handle a large set of quantiles simultaneously, ensuring that it incorporates sufficient information to drive the evolution of the time-varying quantiles.

Yet the main novelty of the DMQ model lies in its use of a score-driven updating scheme. Formally, Catania and Luati (2023) define the updating process as

$$\tilde{q}_t^{\tau_j} = \begin{cases} \tilde{q}_t^{\tau_{j+1}} - \eta_{j,t}, & \text{if } \tau_j < \tau_{j^*} \\ \tilde{q}_t^{\tau_{j^*}}, & \text{if } \tau_j = \tau_{j^*} \\ \tilde{q}_t^{\tau_{j-1}} + \eta_{j,t}, & \text{if } \tau_j > \tau_{j^*} \end{cases} \tag{3.2}$$

with

$$\tilde{q}_t^{\tau_{j^*}} = \tilde{q}_0^{\tau_{j^*}}(1-\beta) + \alpha u_{t-1}^{\tau_{j^*}} + \beta \tilde{q}_{t-1}^{\tau_{j^*}}, \tag{3.3}$$

where $\tilde{q}_0^{\tau_{j^*}}$ denotes the unconditional quantile at the reference probability level, β is the inertia parameter, and α controls the weight of the forcing variable. The process serves as the *reference quantile*, defining the dynamics of the remaining quantiles through

$$\eta_{j,t} = \exp(\xi_{j,t}), \tag{3.4}$$

which are positive-valued random variables with

$$\xi_{j,t} = \xi_{j,0}(1-\phi) + \gamma u_{t-1}^{\tau_j} + \phi \xi_{j,t-1}, \tag{3.5}$$

where ϕ and γ denote the inertia and updating weight parameters, respectively.

The quantile-specific terms $u_t^{\tau_{j^*}}$ and $u_t^{\tau_j}$ are the *forcing variables* responsible for the filtering update defined as

$$u_t^{\tau_j} = \begin{cases} b_j a_j^{-1} \sum_{i=1}^{j} z_{i,t}, & \text{if } j < j^* \\ a_j^{-1} \sum_{i=1}^{j} z_{i,t}, & \text{if } j = j^* \\ b_j a_j^{-1} \sum_{i=j}^{j} z_{i,t}, & \text{if } j > j^* \end{cases} \tag{3.6}$$

where $b_s = \mathbb{1}(\tau_j < \tau_{j^*}) - \mathbb{1}(\tau_j > \tau_{j^*})$ and a_j is a normalizing term defined as in Catania and Luati (2023). This normalization is analogous in spirit to inverse information scaling in score-driven models. However, it is based on the unconditional second moments of empirical quantiles rather than on a parametric likelihood.

The *hit variable* $z_{i,t}$ takes the form of

$$z_t = \frac{\partial}{\partial q_t} \sum_{j=1}^{J} \rho_{\tau_j}\left(y_t - q_t^{\tau_j}\right) = \sum_{j=1}^{J} \frac{\partial}{\partial {q_t}^{\tau_j}} \rho_{\tau_j}(y_t - {q_t}^{\tau_j}) = \mathbb{1}(y_t \leq q_t^{\tau_j}) - \tau_j, \tag{3.7}$$

where $z_t = (z_{i,t},\ i = 1, \ldots, J)'$ and $\rho_\tau(\chi) = \chi\,(\tau - \mathbb{1}(\chi < 0))$ is the quantile check function with the indicator function $\mathbb{1}(\cdot)$.

Importantly, the sequences of the hit variable $\{z_t\}_{t\in\mathbb{N}}$ and the corresponding forcing variable $\{u_t^{\tau_j}\}_{t\in\mathbb{N}}$ are independent and identically distributed (iid) with zero mean and constant variance. Consequently, the updating mechanism of the DMQ model is consistent with the general class of observation-driven models (Creal et al., 2013; Harvey, 2013). Hence, the model provides an observation-driven representation of a predefined set of quantiles within the score-driven framework.

3.2.3 Quantile-heterogeneous and time-varying slopes

The stability of the slope of the Phillips curve remains a topic of ongoing debate. Although some studies find little evidence of significant changes in recent decades (e.g., Hazell et al., 2022), others point to structural changes within the New Keynesian framework, including changes in the Phillips curve itself. For example, Ball and Mazumder (2011) argue that inflation dynamics changed notably after the global financial crisis (GFC), and Rossi et al. (2024) provides evidence of a marked decline in the slope of the Phillips curve since the 1980s. In addition, a complementary finding in Chapter 2 suggests little variation in the slope parameter since the 1960s. However, because of the substantial downside uncertainty in the estimates reported in Figure 2.6, the flattening Phillips curve hypothesis cannot be ruled out. Given these conflicting findings, the question of whether the Phillips curve has genuinely flattened over time remains unresolved.

A natural way to address these considerations within the present framework is to allow the Phillips curve slope parameter to vary across time and quantiles. However, as shown, the score-driven update of time-varying processes requires modeling a large set of quantiles. This poses a nontrivial challenge to the objectives of the study because the original DMQ specification would require a substantial expansion of the parameter space, thereby creating a dimensionality problem. Assigning common parameters to broader quantile ranges mitigates this burden. However, it induces substantial bias by imposing a non-smooth structure on the slope parameters across the distribution, thereby limiting their economic plausibility.

To address this challenge, an alternative approach is adopted to reduce the number of required parameters while retaining sufficient heterogeneity in the slope parameters and enhancing smoothness across quantiles. Specifically, a reduced set of parameter processes, $\hat{\kappa}_{p,t}$, is specified and interpolated to obtain the complete set of quantile-specific parameters at each time t. These time-varying slope parameters are associated with a set of target quantiles defined at predefined probability levels, hereafter referred to as *target probability levels*. Specifically, the analysis considers the standard set $\tau \in \{0.05, 0.25, 0.75, 0.95\}$, which is commonly used in studies of tail risk of macroeconomic variables (e.g., Adrian et al., 2019; López-Salido and Loria, 2024). Additionally, we include the central quantile $\tau = 0.50$, as the median is the standard reference and improves economic interpretation. Empirical evidence in Catania and Luati (2023) further supports this choice, showing that probability levels close to the median minimize the objective function.

The selection of the predefined probability levels serves two primary purposes. First, it ensures representation of the key regions of the distribution while maintaining a balance between parsimony and flexibility. Second, it facilitates identification consistent with theoretical considerations. Under the minimal and economically reasonable assumption that the parameters are monotone between these probability levels, this specification provides sufficient structure to support the interpolation step.

Building on the proposed modifications to the modeling framework, Eq. (3.1) is reformulated using the updated notation as

$$q_t^{\tau_j}(\pi) = \tilde{q}_t^{\tau_j}(\pi) + \exp(\kappa_{j,t})\, x_t + \delta\, \pi_t^E, \tag{3.8}$$

where the full set of conditional inflation quantiles $q_t^{\tau_j}(\pi)$ comprises the directly modeled target quantiles $\hat{q}_t^{\tau_p}(\pi)$ at the target probability levels and auxiliary quantiles $\bar{q}_t^{\tau_j}(\pi)$ at the remaining probability levels. Correspondingly, the full set of

slope parameters $\kappa_{j,t} = (\kappa_{j,t},\ j = 1, \ldots, J)'$ is constructed from the vector of target parameters $\hat{\kappa}_{p,t} = (\hat{\kappa}_{p,t},\ p = 1, \ldots, P)'$, with the remaining elements obtained by interpolation. Specifically, for non-target probability levels, $\kappa_{j,t}$ is deterministically generated from the target parameters $\hat{\kappa}_{p,t}$, which in turn yields the $J - P$ auxiliary quantiles whose slope parameters are not directly estimated.

Formally,

$$\kappa_{j,t} = \mathbb{1}(\tau_j = \tau_p)\, \hat{\kappa}_{p,t} + \mathbb{1}(\tau_j \neq \tau_p)\, \bar{\kappa}_{j,t}, \tag{3.9}$$

where $\bar{\kappa}_{j,t}$ are obtained by linear interpolation within the range spanned by the target parameters and by extrapolation for probability levels below $\tau = 0.05$ and above $\tau = 0.95$. Using this construction, the interpolated parameters are given by

$$\bar{\kappa}_{j,t} = \kappa_t^{\tau_L} + b_j\big(\kappa_t^{\tau_U} - \kappa_t^{\tau_L}\big), \qquad b_j = \frac{\tau_j - \tau_{p_L(j)}}{\tau_{p_U(j)} - \tau_{p_L(j)}}, \tag{3.10}$$

where the $p_L(j) = \arg\max_{p:\, \tau_p < \tau_j}\ \tau_p$ and $p_U(j) = \arg\min_{p:\, \tau_p > \tau_j}\ \tau_p$.

Turning to the specification of the target parameters, their time-varying process follows an autoregressive (AR) process and is defined as

$$\hat{\kappa}_{p,t} = \hat{\kappa}_{p,0} + \phi_p^{\kappa} \hat{\kappa}_{p,t-1} + \rho_p u_{p,t}^{\kappa}, \tag{3.11}$$

where $\hat{\kappa}_{p,0}$ denotes the quantile-specific intercept terms, ϕ_p^{κ} are the inertia parameters, and $u_{p,t}^{\kappa}$ are the forcing variables weighted by ρ_p.

Lastly, the forcing variables are specified as

$$u_{p,t}^{\kappa} = \frac{\partial}{\partial \hat{\kappa}_{p,t}} \sum_{j=1}^{J} \rho_{\tau_j}\big(\pi_t - q_t^{\tau_j}\big) = \sum_{j=1}^{J} \frac{\partial \rho_{\tau_j}}{\partial q_t^{\tau_j}} \frac{\partial q_t^{\tau_j}}{\partial \kappa_{j,t}} \frac{\partial \kappa_{j,t}}{\partial \hat{\kappa}_{p,t}}, \tag{3.12}$$

where the individual components of the forcing variable are obtained as

$$\begin{aligned}
\frac{\partial \rho_{\tau_j}}{\partial q_t^{\tau_j}} &= \mathbb{1}(y_t \leq q_t^{\tau_j}) - \tau_j = z_{j,t}, \\
\frac{\partial q_t^{\tau_j}}{\partial \kappa_{j,t}} &= \frac{\partial}{\partial \kappa_{j,t}}\big(\tilde{q}_t^{\tau_j} + \exp(\kappa_t)\, x_t + \delta\, E_t(\pi_{t+1})\big) = \exp(\kappa_{j,t})\, x_t, \\
\frac{\partial \kappa_{j,t}}{\partial \hat{\kappa}_{p,t}} &= \begin{cases} 1, & \tau_j = \tau_p, \\ (1 - b_j)\, \mathbb{1}(p = p_L(j)) + b_j\, \mathbb{1}(p = p_U(j)), & \tau_j \neq \tau_s. \end{cases}
\end{aligned}$$

The full derivative expression to obtain the forcing variable for updating the time-varying Phillips curve slope parameters is then given by

$$u_{p,t}^{\kappa} = \sum_{j=1}^{J} z_{j,t}\, \exp(\kappa_{j,t})\, x_t \left[(1 - b_j)\, \mathbb{1}\{p = p_L(j)\} + b_j\, \mathbb{1}\{p = p_U(j)\}\right]. \tag{3.13}$$

At this point, it is important to note that constructing the forcing variable in this manner implies that each target TVP depends on the corresponding interpolated parameters. While this design choice confines the information channel to selected regions of the distribution, it also provides a tractable and transparent way to extract information that feeds into the score-driven updates.

Finally, recall that the semi-structural inflation quantiles and their associated parameters at the target probability levels are of primary interest and are modeled directly. The interpolated parameters primarily serve to generate a sufficiently dense set of auxiliary quantiles that, via the forcing variables, provide the information required to update the directly specified target quantiles. Accordingly, the proposed specification should be interpreted as an observation-driven approximation to the evolving conditional inflation distribution rather than as the exact data generating process. The smoothing structure imposed across quantiles therefore serves primarily as a regularization device that facilitates identification and stable estimation. By confining economic interpretation to the target quantiles and using auxiliary quantiles exclusively within the updating scheme, this construction also helps mitigate potential bias arising from the imposed cross-quantile restrictions.

3.2.4 Estimation

The estimation of the model uses the theoretical results of Catania and Luati (2023) for the standard DMQ model. This approach adopts the estimation method first proposed by White et al. (2015).

Specifically, the estimation relies on the quantile check loss function. This function can be interpreted as the negative log-likelihood of an asymmetric Laplace–type density (Koenker and Machado, 1999; Poiraud-Casanova and Thomas-Agnan, 2000; Kotz et al., 2001). More generally, it can be viewed as a member of the tick-exponential family introduced by Komunjer (2005). In the absence of an explicit conditional distributional assumption, minimization of the check loss is equivalent to quasi–maximum likelihood (QML) estimation. Thus, the updating scheme can be viewed as a quasi score-driven (QSD) model, which generalizes the standard score-driven framework outlined in Blasques et al. (2023).

The loss function measures the difference between the observed and predicted values at the probability levels j. Accordingly, the static parameters of the model are estimated by minimizing the objective function of

$$\hat{\theta}_T = \arg\min_{\theta\in\Theta} \sum_{t=1}^{T} \sum_{j=1}^{J} \rho_{\tau_j}\left(y_t - q_t^{\tau_j}(\theta)\right), \tag{3.14}$$

where $q_t^{\tau_j}(\theta)$ denotes the jth conditional inflation quantile, encompassing both target and auxiliary quantile processes. The parameter vector θ collects the static parameters governing the updating dynamics of the J score-driven quantiles and the P time-varying target parameters. The objective function evaluates θ to obtain the parameter values that best fit the observed data.

Given the non-smoothness of the objective function, estimation is performed using Differential Evolution (DE), following Storn and Price (1997), as a global optimization method, followed by local optimization to further improve model fit.

3.3 Benchmark analysis

This section presents the benchmark analysis designed to assess whether the evidence on the Phillips curve relationship obtained in Chapter 2 remains robust within the dynamic multiple quantile Phillips curve model described in Section 3.2. To facilitate a direct comparison across methodologies, the empirical specification is kept as close as possible to the Phillips curve equation introduced in Chapter 2,

Eq. (2.3). In particular, the output gap obtained from the aaGAS model is used as a proxy for real marginal cost, while inflation expectations are assumed to be purely backward-looking. Starting from this parsimonious setup also provides a natural benchmark for the refined specifications introduced in Section 3.4.

3.3.1 Identification and model specifications

Building on the general representation of the dynamic multiple quantile Phillips curve model introduced in Subsection 3.2.1, this section presents its empirical implementation within the sDMQ framework. Relative to Eq. (3.1), the specification incorporates three modifications motivated by identification and estimation considerations. First, the Phillips curve slope parameter is constrained to be positive, consistent with the NKPC theoretical prediction. Second, both the real marginal cost proxy and inflation expectations enter in lagged form to mitigate endogeneity concerns. Third, inflation expectations are proxied by adaptive expectations based on past inflation, consistent with the purely backward-looking NKPC specification considered in Chapter 2.

For ease of reference, the model takes the following form:

$$q_t^{\tau_j}(\pi) = \tilde{q}_t^{\tau_j}(\pi) + \exp(\kappa_{j,t})\, x_{t-1} + \delta\, \pi_{t-1}^E, \tag{3.15}$$

where $\tilde{q}_t^{\tau_j}(\pi)$ denotes the dynamic quantile component defined in Eqs. (3.2)–(3.7). The updating scheme for the time-varying Phillips curve slope $\kappa_{j,t}$ is described in Eqs. (3.9)–(3.13). The marginal cost x_{t-1} is approximated by the real-time estimate of the US output gap $\tilde{y}_{t-1}$ obtained from the aaGAS model introduced in Chapter 2, while π_{t-1}^E denotes adaptive inflation expectations.

As is well known, US inflation exhibits high persistence. This makes the inertia component a key determinant in the present setting. Specifically, the *intrinsic inflation persistence* in the inflation quantile process arises from the AR dependence on past quantiles, $\tilde{q}_{t-1}^{\tau_{j^*}}$ and $\xi_{t-1}^{\tau_j}$ appearing in Eq. (3.3) and Eq. (3.5), respectively. Importantly, by conditioning on information up to time $t-1$, these AR terms substantially reduce the risk of misspecification often arising from omitted lagged inflation dynamics.

The inertia parameters, β, associated with the reference quantile and ϕ, associated with the other quantile processes, are assumed to be time-invariant in the benchmark analysis. This specification preserves some degree of heterogeneity in the persistence dynamics without increasing the model dimension. The choice is supported by the findings of Pivetta and Reis (2007), Benati (2008), and Wolters and Tillmann (2015), who document that although persistence in CPI inflation declined markedly after the early 1980s, persistence in PCE and GDP deflator inflation remained comparatively high and stable.

Higher-order moments of the conditional inflation distribution are of particular interest in the present study. To adequately capture shape and tail behavior, $J = 99$ conditional inflation quantiles are considered. Among these, the $P = 5$ target inflation quantiles at the target probability levels $\tau_p \in \{0.05, 0.25, 0.50, 0.75, 0.95\}$ are directly estimated with their associated semi-structural parameters. The remaining $J - P = 94$ quantiles are treated as auxiliary and primarily serve to inform the score-driven updating mechanism.

Consequently, the full set of dynamic conditional inflation quantiles, including

both target and auxiliary quantiles, is defined as follows:

$$q_t^{\tau_j}(\pi) = \inf\{\, \pi : F_{t|t-1}(\pi) \geq \tau_j \,\}, \quad \tau_j \in \{0.01,\, 0.02,\, \ldots,\, 0.99\}, \quad j = 1, \ldots, J,$$

$$\hat{q}_t^{\tau_p}(\pi) = \inf\{\, \pi : F_{t|t-1}(\pi) \geq \tau_p \,\}, \quad \tau_p \in \{0.05, 0.25, 0.50, 0.75, 0.95\}, \quad p = 1, \ldots, P,$$

$$\bar{q}_t^{\tau_j}(\pi) = \inf\{\, \pi : F_{t|t-1}(\pi) \geq \tau_j \,\}, \quad \tau_j \in \{0.01,\, 0.02,\, \ldots,\, 0.99\}, \quad \tau_j \neq \tau_p,$$

where $F_{t|t-1}(\pi)$ denotes the conditional cumulative distribution function (CDF) of inflation at time t, given information available at time $t-1$.

Several restrictions are imposed to strengthen identification and ensure stable parameter estimation. First, the coefficients governing the score-driven update of the Phillips curve slope parameter, namely the weight of the forcing variable ρ_p and the AR coefficient ψ_p, are restricted to be common across the five target quantiles. In addition, the forcing variables $u_{p,t}^{\kappa}$ are scaled by a normalizing term a_p^{-1}, in a manner similar to Eq. (3.6).

Formally,

$$\bar{u}_{p,t}^{\kappa} = a_p^{-1} u_{p,t}^{\kappa}, \tag{3.16}$$

where $E_{t-1}\left[\left(a_p^{-1} u_{p,t}^{\kappa}\right)^2\right] = 1$ and the scaling constant a_p is defined as in Catania and Luati (2023).

Taking these modifications into account, and after reparameterizing Eq. (3.11), the time-varying Phillips curve slope update is given by

$$\hat{\kappa}_{p,t} = (1 - \phi_\kappa)\hat{\kappa}_{p,0} + \phi_\kappa \hat{\kappa}_{p,t-1} + \rho \bar{u}_{p,t}^{\kappa}. \tag{3.17}$$

The above reparameterization substantially reduces model dimensionality and enhances estimation stability. Importantly, it provides an economically justified approach to setting the initial values of the time-varying parameters, thereby further reducing the parameter space. As a result, the only additional parameters entering the likelihood function are the initial values of the target time-varying parameters $\hat{\kappa}_{p,0}$, along with the common persistence parameter ϕ_κ and the score weight parameter ρ. Economic interpretability also benefits from the new parametrization. When $\phi_\kappa < 1$, $\hat{\kappa}_{p,0}$ represents not merely an initial condition but the long-run reference level, which can be interpreted as the quantile-specific long-run Phillips curve slope.

Several specifications are considered in the the present study. In the empirical analysis, the benchmark DMQ model is estimated and then gradually augmented with semi-structural components, with the final version of the sDMQ NKPC specification with time-varying and quantile-heterogeneous slope parameters. In addition, alternative sDMQ models with static but quantile-specific slope parameters are estimated using three approaches. First, the slope parameters are allowed to vary across predefined ranges of the probability space in a piecewise manner. Second, a monotone piecewise cubic interpolation is employed. Finally, linear interpolation is used. Comparing these methods provides a basis for assessing the benefits of using linear interpolation for the time-varying slope parameters and for identifying potential sources of misspecification.

The fully specified sDMQ model is estimated by jointly estimating the parameter vector

$$\theta = (\alpha, \beta, \gamma, \phi, \delta, \phi_\kappa, \rho, \hat{\kappa}_{p,0})$$

with all parameters constrained to be positive and β, ϕ, ϕ_κ, and δ restricted to lie in the unit interval.

3.3.2 Data description

To analyses the dynamics of US inflation, the seasonally adjusted annualized quarterly growth rate of the core Personal Consumption Expenditures (PCE) price index is used over the period 1961Q1 to 2025Q2. Core PCE inflation serves as the key measure of inflation for the Federal Reserve (Fed), capturing underlying price trends by excluding volatile food and energy components and providing a comprehensive measure of consumer prices that reflects evolving spending patterns.

In line with a purely backward-looking NKPC considered in Chapter 2, adaptive inflation expectations, denoted by π_{t-1}^{E}, are constructed as a moving average of inflation rates realized over the previous three quarters. Given data imprecision, reporting delays, and frequent revisions, the inflation history available at time $t-1$ is used to form expectations. Specifically, adaptive expectations are defined as $\pi_t^E = \frac{1}{3}\sum_{i=1}^{3}\pi_{t-i}$, while long-run persistence is governed separately by the inertia parameters. This specification is consistent with the Phillips curve representation in Chapter 2, Eq. (2.3).

Across all specifications, the output gap, $\tilde{y}_t$, defined as the difference between the logarithm of potential output and the logarithm of real GDP, captures the relationship between inflation and real economic activity, characterized by the time-varying slope of the Phillips curve. The output gap series corresponds to estimates obtained by the aaGAS specification discussed in Chapter 2.

Table 3.2: Summary statistics of PCE inflation rate and output gap (1961Q1–2025Q2)

Variable	$\bar{\mu}$	$\bar{\sigma}$	$\bar{\zeta}$	$\bar{\nu}$	**JB test**
Inflation rate	3.193	2.109	1.215	4.232	79.849 (0.000)
Output gap	0.149	3.021	-0.266	2.794	3.487 (0.175)

Note: JB test refers to the Jarque and Bera test for normality. Parentheses report p-values with significance at the 1% level.

The descriptive statistics of the inflation rate and the output gap over the analyzed period are reported in Table 3.6. As shown, the inflation rate has a positive mean $\bar{\mu}$ of 3.19%, reflecting the long-run average quarterly inflation rate in the sample. Its standard deviation $\bar{\sigma}$ of 2.11 indicates a substantial unconditional volatility.

Positive empirical skewness ($\bar{\zeta} = 1.22$) reveals asymmetry, while excess kurtosis ($\bar{\nu} = 4.23$) suggests a fat-tailed distribution consistent with occasional high-inflation episodes. These characteristics of the empirical distribution suggest that inflation is not normally distributed unconditionally. This is also confirmed by the Jarque and Bera (JB) test, which rejects the null of Gaussianity of the unconditional distribution.

In contrast, the output gap has a mean close to zero ($\bar{\mu} = 0.15$), as expected given its definition as a cyclical deviation from the potential output. Its empirical distribution is slightly negatively skewed ($\bar{\zeta} = -0.27$) and has an excess kurtosis of 2.79. The JB test fails to reject normality, consistent with the moderate skewness and kurtosis values observed for the output gap.

3.3.3 Empirical comparison of DMQ and sDMQ specifications

In this section, the baseline DMQ model with and without inflation determinants are compared with the sDMQ NKPC model with backward-looking inflation expectations across alternative specifications.

Three methods are considered for constructing the auxiliary parameters in the static-parameter sDMQ model. First, parameters are held constant across subsets of probability levels, yielding piecewise relationships across the corresponding regions of the probability space. Specifically, the quantile indices $j = 1, \ldots, 99$ are partitioned into the ranges $j \in \{1:9,\ 10:39,\ 40:60,\ 61:90,\ 91:99\}$, and a common slope parameter κ_p is estimated for all quantiles within each range. Although this approach is strongly restrictive and lacks economic justification, it serves as a useful benchmark for comparison. The second and the third specifications use a monotone piecewise cubic interpolation (MPCI) described in Fritsch and Carlson (1980) and linear interpolation (LI), respectively. Finally, the sDMQ NKPC model with time-varying and quantile-heterogeneous slope parameters is considered.

For clarity and ease of navigation, Table 3.3 provides an overview of the model specifications considered in this section and their shorthand notation. Parameter estimates obtained from the evaluated models, together with the corresponding objective function values, are reported in Table 3.4.

Table 3.3: Summary of model specifications: Benchmark analysis

Model	Description
DMQ(1)	Benchmark DMQ model of Catania and Luati (2023)
DMQ(2)	DMQ model with static and quantile-invariant parameters
sDMQ(1)	sDMQ model with piecewise quantile range-specific PC slope
sDMQ(2)	sDMQ NKPC with MPCI PC slope
sDMQ(3)	sDMQ NKPC with LI PC slope
sDMQ-TV	sDMQ NKPC with time-varying and quantile-specific PC slope

Improvements in model fit already emerge when comparing the benchmark DMQ(1) model with its semi-structural augmented counterpart, DMQ(2). While parameter values associated with the score-driven update of the dynamic quantiles are broadly comparable across these specifications, the somewhat lower estimates of α and β suggest that the inclusion of inflation determinants absorbs part of the information relevant for updating the dynamic inflation quantiles.

The gain in model fit becomes evident once the assumption of a quantile-homogeneous Phillips curve slope is relaxed, even after accounting for the increase in the number of parameters. However, the magnitude of this gain varies substantially across the approaches used to construct quantile-specific parameters in the sDMQ models. As expected, the piecewise specification delivers the weakest fit in terms of the aggregate loss, reflecting the strong economic restrictions imposed by abrupt parameter changes across quantiles. In contrast, interpolation-based approaches yield substantial improvements in model fit. Although the lowest loss value is achieved under the MPCI specification, linear interpolation provides a comparable fit, indicating a reasonable trade-off between parsimony and flexibility.

Figure 3.1 provides a visual overview of the estimated slope parameter and shows that its overall shape is broadly comparable across the three methods. Most importantly, the comparison reveals pronounced quantile heterogeneity in the PC slope. The slope reaches its highest values at lower-middle quantiles of the inflation

Table 3.4: Parameter estimates across model specifications (1961Q1-2025Q2)

	DMQ(1)	**DMQ(2)**	**sDMQ(1)**	**sDMQ(2)**	**sDMQ(3)**	**sDMQ-TV**
Score-driven update parameters						
α	0.527	0.490	0.474	0.483	0.469	0.466
β	1.000	0.997	0.987	0.977	0.971	0.993
γ	0.072	0.072	0.063	0.096	0.039	0.101
ϕ	0.998	0.998	0.998	0.997	0.998	0.997
ρ	–	–	–	–	–	0.2035
ϕ_κ	–	–	–	–	–	0.9998
Semi-structural parameters						
δ	–	0.079	0.141	0.227	0.242	0.172
κ	–	0.005	0.039	0.037	0.048	0.036
$\kappa_{\tau=0.05}$	–	–	0.052	0.024	0.049	0.037 (0.009)
$\kappa_{\tau=0.25}$	–	–	0.095	0.107	0.137	0.200 (0.078)
$\kappa_{\tau=0.50}$	–	–	0.030	0.049	0.046	0.096 (0.061)
$\kappa_{\tau=0.75}$	–	–	0.003	0.003	0.004	0.030 (0.016)
$\kappa_{\tau=0.95}$	–	–	0.014	0.003	0.003	0.097 (0.017)
Loglik.	6021.80	6011.59	5965.99	5949.29	5955.12	5940.44
AIC	12051.60	12035.18	11943.98	11910.58	11922.24	11904.84

Note: Values of κ under the sDMQ specification are averages across quantiles, while under sDMQ-TV they are averaged across both time and quantiles. Parameter estimates of the PC slope under sDMQ-TV correspond to $\hat{\kappa}_{p,0}$, with sample averages reported in parentheses. The reported Loglik. values correspond to the final value of negative quasi-maximum likelihood (QML) objective function evaluated at the estimated parameter values.

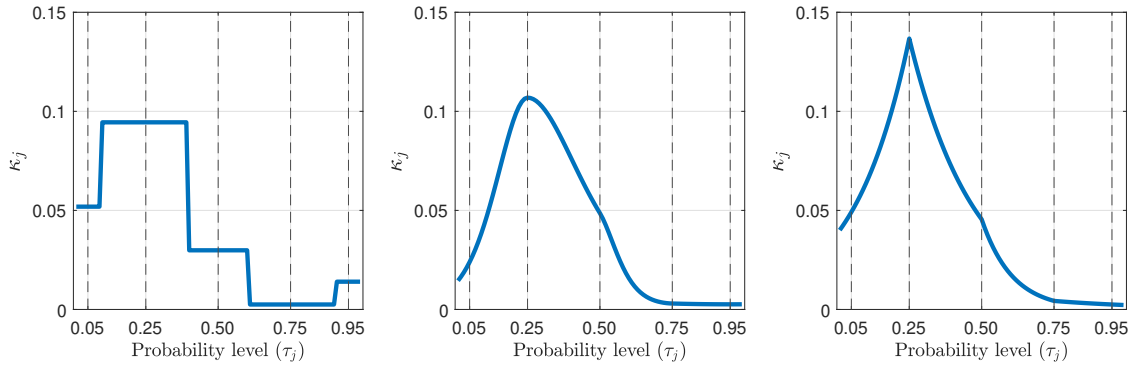

Figure 3.1: Phillips curve slope across inflation quantiles. The left panel reports estimates obtained using the piecewise method sDMQ(1), the middle panel shows results based on monotone piecewise cubic interpolation sDMQ(2), and the right panel displays estimates based on linear interpolation sDMQ(3). Vertical dashed lines indicate the target probability levels.

distribution, with a peak around the first quartile. This pattern suggests that inflation reacts more strongly to economic conditions in case of downside risks, but not in the extreme lower tail. As one moves toward the median and into higher quantiles, the slope declines markedly, suggesting a progressively flatter slope at higher inflation outcomes.

Overall, these results imply that the trade-off between inflation and the output gap varies substantially across the inflation distribution, underscoring the importance of allowing for quantile-specific slope parameters rather than imposing homogeneity.

It is worth noting that, because κ_p is modeled in logarithmic form, interpolation is performed on the untransformed values. This leads to smoother transitions between the target parameters and, in this setting, yields a slope profile with greater

economic plausibility. Therefore, the implied curvature between target quantiles should be interpreted as a reduced-form feature of the interpolation scheme rather than as a structural restriction.

When comparing all considered models, the inertia parameter β is estimated at high levels consistent with the intrinsic persistence of PCE inflation. However, their values decline modestly under more flexible specifications, suggesting that the semi-structural components absorb part of the inertia. A similar pattern emerges for the weight parameter α, indicating that inflation determinants contain relevant information for updating the dynamic quantiles.

Turning to the sDMQ-TV model, the time- and quantile-averaged value of the time-varying slope parameter $\kappa_{p,t}$ is comparable across the sDMQ specifications and consistent with the sample average of the estimated time-varying Phillips curve slope parameter, $b_\pi = 0.04$, reported in Chapter 2 (Table 2.3). While the quantile-specific sample averages display a shape similar to their static counterparts, they differ somewhat in magnitude. Interestingly, the average parameter values are closest to those obtained under the sDMQ(2) specification, making the implied average slope profile under sDMQ-TV most similar to that specification. Importantly, the sDMQ-TV specification yields the lowest aggregate loss values and the lowest Akaike information criterion (AIC), indicating the best overall model fit.

In summary, the results indicate pronounced heterogeneity in the trade-off between inflation and the output gap across the distribution of inflation. Comparisons between alternative methods for capturing quantile heterogeneity in the slope consistently point to an asymmetric hump-shaped slope profile. Parameter estimates of $\kappa_{p,0}$ and their associated sample averages suggest that the hump-shaped profile is preserved when the slope is allowed to vary over time. Notably, model fit improves substantially under the sDMQ specifications relative to the benchmark DMQs, with the sDMQ-TV specification delivering the best overall model fit.

3.3.4 The historical evolution of the Phillips curve slope

Figure 3.2 provides a historical overview of the time-varying estimates of the Phillips curve slope parameter across the full set of inflation quantiles, associated with both target and auxiliary probability levels.

The estimated Phillips curve slope exhibits pronounced regime-dependent patterns that align closely with major macroeconomic episodes in the postwar US economy. Overall, the figure reveals substantial heterogeneity in the Phillips curve slope, with an asymmetric hump-shaped profile across the distribution. This distributional asymmetry is particularly pronounced during the 1970s, when the slope is markedly steeper and exhibits greater variability than in the post-1985 period.

Following the Volcker disinflation, the slope declines sharply, with the most pronounced reduction observed around the median and the first quartile of the distribution. This decline reflects a substantial weakening in the inflation–output gap trade-off, likely induced by the aggressive tightening of monetary policy and the subsequent anchoring of inflation expectations. The slope remains relatively flat at the tails, implying that extreme inflation and deflationary outcomes are weakly responsive to real economic activity and are instead driven by expectations and persistence. While the slope around the median remains largely intact in the post-COVID period, it further flattens at the lower-central region of the distribution.

To facilitate quantitative assessment, Figure 3.3 restricts attention to the slope parameters at the target probability levels. Consistent with the discussion above,

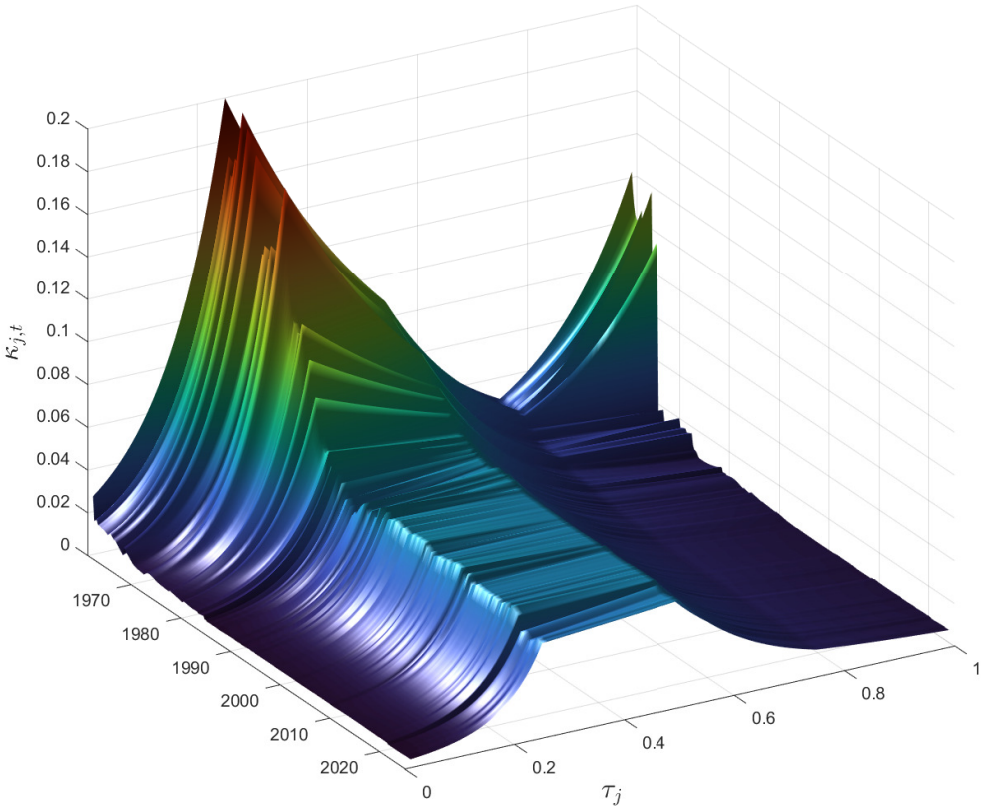

Figure 3.2: Historical time-varying NKPC slope parameters across the US conditional inflation distribution, based on estimates from the sDMQ-TV model.

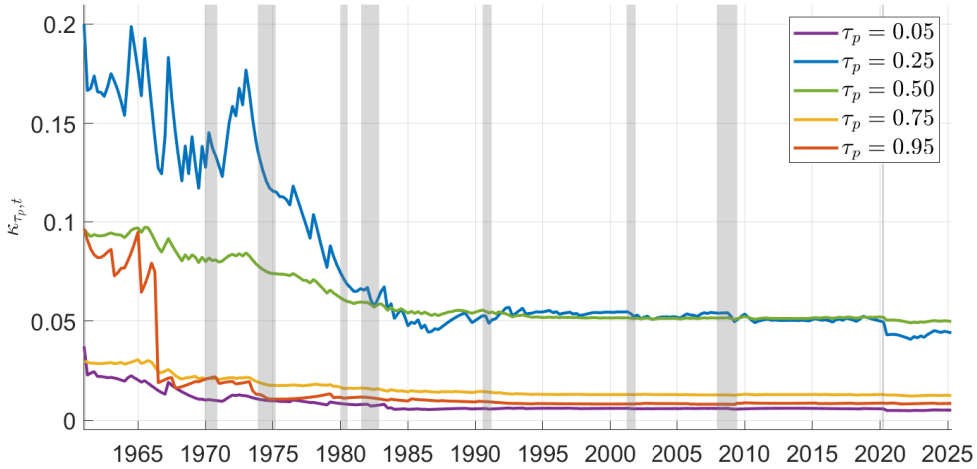

Figure 3.3: Historical time-varying NKPC slope parameters at target quantiles, based on estimates from the sDMQ-TV model.

the figure shows a pronounced decline in the Phillips curve slope across all quantiles since the 1960s, with the largest reductions occurring at the $\tau = 0.25$ probability level. In general, the slope parameters remain relatively stable from the mid-1980s onward, with modest variability at the first quartile during periods of macroeconomic stress. Notably, the estimated slope at the median is around 5% in the

post-1985 period, broadly consistent with the median of the simulation output reported in Chapter 2 (Figure 2.6, right-hand panel). In addition, the associated downside uncertainty shown in Figure 2.6 is consistent with the asymmetric hump-shaped Phillips curve slope across quantiles, characterized by relatively steeper slopes at lower quantiles and flatter slopes at upper quantiles.

Overall, these results highlight the importance of allowing for both time variation and distributional heterogeneity in the historical analysis of US inflation dynamics. The sDMQ-TV specification captures economically meaningful regime shifts that are obscured in models imposing constant or quantile-invariant slopes. Importantly, the results suggest that the conflicting empirical evidence on the Phillips curve flattening hypothesis, as well as the substantial downside uncertainty surrounding the slope estimates obtained in Chapter 2, are likely driven by pronounced slope heterogeneity and asymmetry across the inflation distribution.

At the same time, the benchmark analysis remains subject to several limitations. In particular, the reliance on adaptive inflation expectations constrains the economic interpretation of the estimated Phillips curve dynamics and limits consistency with the standard NKPC theoretical framework. More importantly, the use of adaptive expectations based on past inflation makes it difficult to disentangle the effects of inflation expectations and inflation inertia, since both channels are driven by past inflation and may therefore be highly correlated. This complicates the identification of their respective contributions to inflation dynamics. These limitations motivate the refined empirical analysis presented in the following section.

3.4 Refined model specification and main results

While the benchmark empirical analysis in Section 3.3 provides evidence of pronounced heterogeneity in the Phillips curve slope profile, likely contributing to the documented downside uncertainty surrounding its estimate in Chapter 2, its exclusive reliance on past inflation to form inflation expectations limits its consistency with the New Keynesian theoretical framework. In addition, assuming purely adaptive price setting behavior prevents an assessment of the relative roles of backward- and forward-looking inflation expectations, which are central to macroeconomic policy discussions. Therefore, the present section considers a hybrid NKPC specification with a time-varying relative weight on backward- and forward-looking inflation expectations.

In addition, the main analysis addresses several limitations of the benchmark analysis through a sequence of extensions gradually introduced in the present section. First, Subsection 3.4.1 incorporates forward-looking inflation expectations into the model, leading to a key modification of the original DMQ updating process proposed by Catania and Luati (2023) and applied in Section 3.3. Second, Subsection 3.4.2 integrates monotone piecewise interpolation into the time-varying updating mechanism of the quantile-heterogeneous sDMQ-TV framework. Third, Subsection 3.4.3 introduces additional constraints that improve the identification of the relative role of inflation expectations and enhance estimation stability. Fourth, Subsection 3.4.4 describes the construction of short- and long-run inflation expectation measures and introduces a new proxy for real marginal cost. Finally, the remaining subsections present the main empirical results, including parameter estimates in Subsection 3.4.5, Phillips curve slope dynamics in Subsection 3.4.6, the relative role of inflation expectations in Subsection 3.4.7, and the evolution of

the conditional inflation quantiles and conditional moments in Subsections 3.4.8 and 3.4.9, respectively.

3.4.1 Inflation inertia and expectations dynamics

The key departure from the benchmark model employed in Section 3.3 in the present analysis is to allow the inertia parameter β associated with the dynamics of the reference quantile to vary over time and to introduce survey-based inflation expectations to capture forward-looking inflation expectations. To this end, the inflation expectation component in Eq. (3.1) is incorporated directly into the reference quantile process by replacing the unconditional reference quantile level $\tilde{q}_0^{\tau_j*}$ in Eq. (3.3). This modification provides an economically interpretable anchoring mechanism for the conditional inflation distribution while preserving the parsimonious structure of the original DMQ framework and avoiding an expansion of the parameter space.

Importantly, this specification differs fundamentally from the benchmark analysis, which relies exclusively on adaptive inflation expectations. By replacing the unconditional reference quantile with survey-based inflation expectations, the model transitions from a purely backward-looking specification to a hybrid NKPC framework. This modification substantially alters the identification and interpretation of the persistence and expectations channels within the model, making direct comparison with the benchmark specification less straightforward and warranting separate treatment in the main empirical analysis.

This representation is also consistent with a commonly imposed restriction in the hybrid NKPC literature where inflation coefficients sum to one (e.g., Mavroeidis et al., 2014). Such a restriction facilitates the interpretation of inflation dynamics in terms of the relative importance of expectations and inertia. In the present framework, the relative importance of these two components is governed by the time-varying parameter β_t, allowing the contribution of forward- and backward-looking expectations to evolve over time.

Accordingly, the reference quantile process in Eq. (3.3) is redefined as

$$\tilde{q}_t^{\tau_j*} = (1-\beta_t)\pi_{t-1}^E + \alpha u_{t-1}^{\tau_j*} + \beta_t \tilde{q}_{t-1}^{\tau_j*}. \tag{3.18}$$

The updating equation for the time-varying weight parameter takes the form

$$\beta_t = (1-\phi_\beta)\beta_0 + \phi_\beta \beta_{t-1} + \rho_\beta u_t^\beta, \tag{3.19}$$

where ϕ_β denotes the inertia parameter, ρ_β the loading coefficient on the forcing variable u_t^β, and β_0 the long-run reference weight, capturing the relative importance of backward- and forward-looking components. During estimation, a logistic link function is used to ensure that $0 < \beta_t < 1$, allowing β_t to be interpreted as a time-varying weight governing the state-dependent relative importance of inflation expectations and inflation inertia.

Lastly, the forcing variable u_t^β is constructed from the score of the quantile loss function and enters the updating equation through a quasi-score specification:

$$u_t^\beta = z_t \, \beta_t \left(\tilde{q}_{t-1}^{\tau_j*} - \pi_{t-1}^E \right). \tag{3.20}$$

It should be noted that, by construction, beyond its direct effect on the reference quantile, inflation expectations affect the dynamics of the remaining quantile processes indirectly through the reference quantile, in a manner analogous to the inertia parameter β in the baseline DMQ specification.

3.4.2 Improved interpolation scheme

Beyond the economic considerations motivating the hybrid NKPC specification, Table 3.4 suggests that the baseline empirical specification also involves methodological trade-offs, reflected in the weaker model fit obtained under specifications relying on linear interpolation. While linear interpolation offers a parsimonious closed-form construction, it may introduce misspecification at auxiliary quantiles not directly linked to the target TVPs. Although careful selection of target probability levels can mitigate this issue, it cannot eliminate it entirely. Accordingly, this section applies monotone piecewise cubic interpolation consistently across all specifications.

Building on the proposed modifications to the modeling framework described in Subsection 3.4.1, Eq. (3.1) is reformulated as

$$q_t^{\tau_j}(\pi) = \tilde{q}_t^{\tau_j}(\pi) + \exp(\kappa_{j,t})\, x_t, \tag{3.21}$$

where the notation follows the baseline specification with the updated interpolation scheme. Formally,

$$\kappa_{j,t} = \mathbb{1}(\tau_j = \tau_p)\, \hat{\kappa}_{p,t} + \mathbb{1}(\tau_j \neq \tau_p)\, \bar{\kappa}_{j,t}, \tag{3.22}$$

where $\bar{\kappa}_{j,t}$ are now obtained by monotone piecewise cubic interpolation (MPCI), as described in Fritsch and Carlson (1980), within the range spanned by the target parameters, and by extrapolation for probability levels below $\tau = 0.05$ and above $\tau = 0.95$. Using this construction, the interpolated parameters are given by

$$\bar{\kappa}_{j,t} = h_{00}(b_j)\, \kappa_t^{\tau_L} + h_{01}(b_j)\, \kappa_t^{\tau_U}, \qquad b_j = \frac{\tau_j - \tau_{p_L(j)}}{\tau_{p_U(j)} - \tau_{p_L(j)}}, \tag{3.23}$$

where $p_L(j) = \arg\max_{p:\, \tau_p < \tau_j}\ \tau_p$ and $p_U(j) = \arg\min_{p:\, \tau_p > \tau_j}\ \tau_p$, and the MPCI basis functions are

$$h_{00}(b) = 2b^3 - 3b^2 + 1, \qquad h_{01}(b) = -2b^3 + 3b^2. \tag{3.24}$$

Turning to the specification of the target parameters, their time-varying process follows an autoregressive (AR) process and is defined as

$$\hat{\kappa}_{p,t} = \hat{\kappa}_{p,0} + \phi_p^{\kappa} \hat{\kappa}_{p,t-1} + \rho_p u_{p,t}^{\kappa}, \tag{3.25}$$

where $\hat{\kappa}_{p,0}$ denotes the quantile-specific intercept terms, ϕ_p^{κ} are the inertia parameters, and $u_{p,t}^{\kappa}$ are the forcing variables weighted by ρ_p.

Lastly, the forcing variables are specified as

$$u_{p,t}^{\kappa} = \frac{\partial}{\partial \hat{\kappa}_{p,t}} \sum_{j=1}^{J} \rho_{\tau_j}\big(\pi_t - q_t^{\tau_j}\big) = \sum_{j=1}^{J} \frac{\partial \rho_{\tau_j}}{\partial q_t^{\tau_j}} \frac{\partial q_t^{\tau_j}}{\partial \kappa_{j,t}} \frac{\partial \kappa_{j,t}}{\partial \hat{\kappa}_{p,t}}, \tag{3.26}$$

where the individual components of the forcing variable are obtained as

$$\frac{\partial \rho_{\tau_j}}{\partial q_t^{\tau_j}} = \mathbb{1}(\pi_t \leq q_t^{\tau_j}) - \tau_j = z_{j,t},$$

$$\frac{\partial q_t^{\tau_j}}{\partial \kappa_{j,t}} = \exp(\kappa_{j,t})\, x_t,$$

$$\frac{\partial \kappa_{j,t}}{\partial \hat{\kappa}_{p,t}} = \begin{cases} 1, & \tau_j = \tau_p, \\ h_{00}(b_j)\, \mathbb{1}\{p = p_L(j)\} + h_{01}(b_j)\, \mathbb{1}\{p = p_U(j)\}, & \tau_j \neq \tau_p. \end{cases}$$

The resulting forcing variable is therefore

$$u_{p,t}^{\kappa} = \sum_{j=1}^{J} z_{j,t} \exp(\kappa_{j,t})\, x_t \left[h_{00}(b_j)\, \mathbb{1}\{p = p_L(j)\} + h_{01}(b_j)\, \mathbb{1}\{p = p_U(j)\} \right]. \quad (3.27)$$

3.4.3 Main empirical specifications

Building on the framework presented in Subsections 3.4.1–3.4.2, we consider the sDMQ hybrid NKPC model

$$q_t^{\tau_j}(\pi) = \tilde{q}_t^{\tau_j}(\pi) + \exp(\kappa_{j,t})\, x_{t-1}, \quad (3.28)$$

where $\tilde{q}_t^{\tau_j}(\pi)$ denotes the dynamic quantile component defined in Eqs. (3.2)–(3.7), as in the baseline model, but with the modified reference quantile process introduced in Eqs. (3.18)–(3.20). The updating scheme with piecewise monotone interpolation for the time-varying Phillips curve slope $\kappa_{j,t}$ is described in Eqs. (3.22)–(3.27).

Beyond estimating the fully specified sDMQ-TV model defined in Eq. (3.28), the main empirical analysis considers a sequence of additional specifications starting from the baseline DMQ model, which abstracts from structural relationships. In addition to enhancing estimation stability and enabling the evaluation of incremental model improvements, this approach facilitates a clearer interpretation of the contribution of each modeling component. For clarity, Table 3.5 summarizes the model specifications considered in this section and their shorthand notation.

Table 3.5: Summary of model specifications: Main analysis

Model	Description
DMQ	Baseline DMQ model with the specification of Catania and Luati (2023)
DMQ(S/L)	DMQ NKPC with static quantile-homogeneous parameters (β, κ)
DMQ-TV(S/L)	DMQ model with time-varying quantile-homogeneous parameters (β_t, κ_t)
sDMQ(S/L)	sDMQ NKPC with static quantile-heterogeneous slopes (β, κ_p)
sDMQ-TV(S/L)	sDMQ NKPC with time-varying quantile-heterogeneous slopes (β_t, $\kappa_{p,t}$)

Note: S and L denote models with short-run and long-run inflation expectations, respectively.

Augmenting the baseline DMQ model with time-invariant inflation determinants represents a useful intermediate step. First, it provides a basis for evaluating the relevance of moving from a fully reduced-form inflation dynamics toward a more structural specification. Second, the time-invariant estimation of the relative weights of inflation inertia and forward-looking expectations β and the PC slope κ provides a useful anchor for the time-varying specifications (DMQ-TV). Specifically, the parameterization of the time-varying evolution of these parameters, as in Eq. (3.18), allows these estimates to serve as long-run reference values, thereby improving identification. Similarly, by capturing quantile heterogeneity in the Phillips curve slope, the time-invariant sDMQ model provides a natural basis for treating these estimates as long-run reference values in the sDMQ-TV specification.

While the DMQ-TV specification helps assess time variation, the sDMQ model with static parameters allows for the assessment of quantile heterogeneity. Comparing the estimation results of these specifications allows the two dimensions to be assessed in isolation, whereas the sDMQ-TV model enables their joint assessment.

Several additional restrictions are imposed in the main analysis to strengthen identification and ensure stable estimation of the parameters in the final sDMQ-TV

model. Specifically, besides forcing the weight parameters in the updating process, ρ_p, and the AR coefficient, ϕ_p^κ, to be common across the five target quantiles, as in the benchmark analysis, the initial values and long-run reference values are fixed at the corresponding time-invariant estimates obtained from the sDMQ model.

Taking these constrains into account, and after reparameterizing Eq. (3.25), the time-varying Phillips curve slope update is given by

$$\hat{\kappa}_{p,t} = (1 - \phi_\kappa)\hat{\kappa}_{p,0} + \phi_\kappa \hat{\kappa}_{p,t-1} + \rho u_{p,t}^\kappa. \tag{3.29}$$

This reparameterization reduces model dimensionality and enhances estimation stability. Importantly, it provides an economically justified approach to setting the initial values of the time-varying parameters, thereby further reducing the effective parameter space.

Given the relative sparsity of inflation realizations in the tails and the numerical instability induced by interpolation, which can exacerbate the non-smoothness of the likelihood function in the DMQ framework, the static slope parameters may be weakly identified. We address this issue along several dimensions. First, rather than relying on extrapolation beyond the extreme probability levels, we fix the slope in the tail regions at the value estimated at the corresponding boundary quantiles. Second, we impose a lower bound on the slope parameters, $\kappa_p \geq 0.0025$, to enhance numerical stability.

Finally, instead of fixing the initial values $\hat{\kappa}_{p,0}$ to the time-invariant sDMQ estimates, we adopt a shrinkage-based parameterization that allows for partial pooling across quantiles. Specifically, $\hat{\kappa}_{p,0}$ is modeled as a convex combination of the quantile-specific estimates $\hat{\kappa}_p$ obtained by the sDMQ specification and a homogeneous long-run reference $\bar{\kappa}$, defined as the cross-quantile average of the five target quantile estimates, with the shrinkage intensity governed by the parameter ψ_s.

Formally,

$$\hat{\kappa}_{p,0} = \bar{\kappa} + \psi_s \left(\hat{\kappa}_p - \bar{\kappa}\right), \tag{3.30}$$

where $\psi_s \in [0, 1]$ determines the degree of slope heterogeneity across quantiles: the closer its value is to unity, the closer the heterogeneity is to that estimated in the static sDMQ specification.

An important aspect of the DMQ framework is the normalization of the forcing variables across quantile processes. Because the construction of the forcing variables $u_{p,t}^\kappa$ defined in Eq. (3.26) relies on multiple components, scaling by a single normalizing term a_p^{-1}, as in Eq. (3.6), is no longer appropriate, as it fails to account for the parameter dependence of the forcing variables.

To address this problem, we adopt a data-driven normalization strategy that accounts for the dependence of the variance of the forcing variables on the shrinkage parameter. Specifically, we precompute a variance surface $\hat{V}_p(\psi_s; \hat{\kappa}_p)$ over a grid $\psi_s \in \{0, 0.1, \ldots, 1\}$, where the variance is evaluated under a fixed-parameter specification, i.e., $\rho = 0$. During estimation, this surface is evaluated at the current value of ψ_s via shape-preserving spline interpolation, yielding $\bar{V}_p(\psi_s; \hat{\kappa}_p) = \mathcal{P}(\psi_s; \hat{V}_p)$. The forcing variables are then normalized as $\bar{u}_{p,t}^\kappa = \bar{V}_p(\psi_s; \hat{\kappa}_p)^{-1/2} u_{p,t}^\kappa$. This procedure ensures that the scaling of the forcing variables appropriately reflects the underlying parameterization, leading to comparable variances across quantiles, which stabilizes the dynamics and improves optimization behavior.

Taken together, the fully specified sDMQ-TV model is estimated by jointly estimating the following parameter vector, with β_0 and $\hat{\kappa}_{p,0}$ obtained from the

sDMQ model and ϕ_β from the DMQ-TV model:

$$\theta = (\alpha, \gamma, \phi, \phi_\kappa, \rho_\kappa, \rho_\beta, \psi_s),$$

where all parameters are constrained to be positive, and α, ϕ, and ϕ_κ are further restricted to lie within the unit interval.

3.4.4 Data description

The present analysis retains the seasonally adjusted annualized quarterly growth rate of the core Personal Consumption Expenditures (PCE) price index employed in the benchmark analysis (Section 3.3), while extending the sample period to cover 1961Q1–2025Q4.

In contrast to the benchmark analysis, which relied on filtered output gap estimates obtained from the aaGAS model, the empirical specification considered here employs the output gap series published by the Congressional Budget Office (CBO). This choice is motivated by the structural nature of the CBO measure, where potential output is derived from a production function framework incorporating trend labor input, capital services, and total factor productivity. In addition, the CBO output gap is widely used in empirical macroeconomic analysis, thereby facilitating comparability with the existing literature.

We proxy inflation expectations using the 2-year inflation expectations series from the Federal Reserve Bank of Cleveland, obtained from FRED at a quarterly frequency and available from 1982Q2 onward. Following the approach of Benigno and Eggertsson (2023), the series is extended backward to 1961Q1 using 12-month inflation expectations from the Livingston Survey. As the latter is available at a semiannual frequency, missing quarterly observations are interpolated using a spline-based, curve-preserving method.

In addition, we use long-run PCE inflation expectations as an alternative longer-horizon measure of expected inflation to assess the sensitivity of the results to the expectations specification and facilitate comparison across alternative expectation measures. Following the approach of Clark and Doh (2014), also adopted by Del Negro et al. (2017), the series is constructed by combining Survey of Professional Forecasters (SPF) data from 2007Q1 onward with survey-based long-run (5- to 10-year-ahead) PCE inflation expectations from the Federal Reserve Board's FRB/US econometric model for earlier periods.

Figure 3.4 displays the inflation series and Phillips curve determinants used in the analysis. The upper panel shows US core PCE inflation alongside short-run and long-run inflation expectations over the sample period. Core PCE inflation is highly volatile during the 1970s and early 1980s, followed by a sustained moderation from the mid-1980s, and a renewed increase in the post-COVID period. Long-run inflation expectations co-move with PCE inflation but exhibit smoother dynamics and attenuated short-run fluctuations. In contrast, short-run expectations display greater variability and respond more strongly to cyclical movements in inflation.

The lower panel reports the CBO output gap over the sample period. The series exhibits pronounced cyclical dynamics, with large negative gaps during major recessions, notably in the early 1980s, the Great Recession, and the COVID-19 shock in 2020. Overall, the figure highlights the joint evolution of inflation, expectations, and real economic slack, which motivates the empirical Phillips curve specification.

The descriptive statistics of the inflation rate, the output gap, and inflation expectations over the analyzed period are reported in Table 3.6. As shown, the inflation rate has a positive mean $\bar{\mu}$ of 3.19%, reflecting the long-run average quarterly

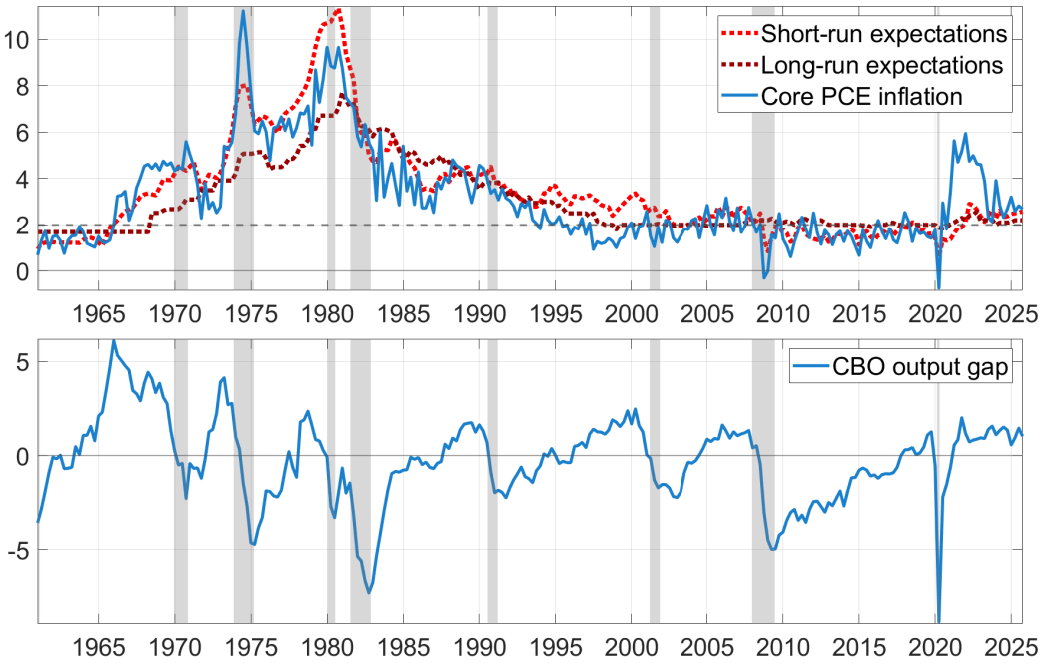

Figure 3.4: US core PCE inflation and inflation expectations (upper panel) and CBO output gap (lower panel), 1961Q1–2025Q4.

Table 3.6: Summary statistics of inflation and its determinants (1961Q1–2025Q4)

Variable	$\bar{\mu}$	$\bar{\sigma}$	$\bar{\zeta}$	$\bar{\nu}$	**JB test**
Inflation rate	3.192	2.099	1.224	4.274	82.498 (0.000)
Short-run expectations	3.446	2.190	1.542	5.234	157.160 (0.000)
Long-run expectations	3.052	1.508	1.203	3.327	63.840 (0.000)
Output gap	-0.248	2.254	-0.381	4.086	19.072 (0.000)

Note: JB test refers to the Jarque and Bera test for normality. Parentheses report p-values. All series reject normality at the 1% level.

inflation rate in the sample. Its standard deviation $\bar{\sigma}$ of 2.10 indicates substantial unconditional volatility. Positive empirical skewness ($\bar{\zeta} = 1.22$) reveals asymmetry, while kurtosis ($\bar{\nu} = 4.27$) suggests a fat-tailed distribution consistent with occasional high-inflation episodes.

In contrast, the output gap has a mean close to zero ($\bar{\mu} = -0.25$), as expected given its definition as a cyclical deviation from potential output. Its empirical distribution is slightly negatively skewed ($\bar{\zeta} = -0.38$) and exhibits elevated kurtosis ($\bar{\nu} = 4.09$), indicating somewhat heavier tails than a Gaussian benchmark.

Turning to inflation expectations, short-run expectations display a higher mean ($\bar{\mu} = 3.45$) and greater dispersion ($\bar{\sigma} = 2.19$) than realized inflation. Their distribution is strongly right-skewed ($\bar{\zeta} = 1.54$) and leptokurtic ($\bar{\nu} = 5.23$), pointing to pronounced asymmetry and tail risk. By contrast, long-run expectations are more tightly anchored, with a lower standard deviation ($\bar{\sigma} = 1.51$) and kurtosis ($\bar{\nu} = 3.33$), and a mean ($\bar{\mu} = 3.05$) close to that of realized inflation. Their skewness ($\bar{\zeta} = 1.20$) remains positive but less pronounced than for short-run expectations.

The characteristics of the empirical distributions suggest that the observables are not normally distributed unconditionally. This is confirmed by the Jarque–Bera (JB) test, which rejects the null hypothesis of Gaussianity for each series.

3.4.5 Parameter estimates

This section reports the estimation results for the specifications summarized in Table 3.5 using short-run (S) and long-run (L) inflation expectations.

An important consideration in the use of survey-based inflation expectations concerns their measurement and informational content. As emphasized by Mavroeidis et al. (2014), survey expectations are potentially contaminated by a composite error term reflecting both measurement error and information frictions, arising from the reliance of survey respondents on a more limited information set than that assumed in the model, as well as their limited ability to filter out cost-push shocks. As a result, survey-based expectations are subject to noise and may not be strictly exogenous.

While lagging the expectations measure helps alleviate endogeneity concerns, it does not resolve the fundamental identification problem that arises when expectations exhibit limited variation. In such cases, their contribution to inflation dynamics becomes weakly identified, as emphasized by Mavroeidis et al. (2014) and discussed more broadly in Cochrane (2011). This implies that the empirical role of expectations may depend critically on the variability of the expectation proxy employed. In this context, alternative measures that are more strongly anchored, such as long-run expectations, may yield substantially different implications for both the estimated weight on expectations and the implied dynamics of the Phillips curve.

The estimated parameters and corresponding objective function values are reported in Tables 3.7 and 3.8. The aggregate loss and information criteria in Table 3.7 indicate that model fit improves substantially when moving from the baseline DMQ model to its PC-augmented counterparts, DMQ(S) and DMQ(L). The gain in model fit is more pronounced under the DMQ(S) specification, which is associated with steeper estimated slope parameter and larger expectation weights relative to the corresponding specification based on long-run expectations. This suggests that short-run inflation expectations provide stronger identifying variation in the estimated models, consistent with the identification issues discussed above. Whether the nearly flat slope parameters in the homogeneous DMQ(L) and DMQ-TV(L) specifications reflect weak identifying variation associated with long-run inflation expectations warrants further investigation.

Parameters governing the score-driven updating of the dynamic quantiles are broadly comparable across specifications, with some notable exceptions. First, the estimate of β under the benchmark DMQ model is close to unity ($\beta = 0.99999$), suggesting that the unconditional median contributes little information to the dynamic quantile updating process. In contrast, the lower estimates of β under the semi-structural specifications suggest that incorporating inflation expectations into the reference quantile process reduces the inertia, and hence the role of the backward-looking component. Second, when comparing specifications based on the two types of expectations, the higher estimates of β and α associated with long-run expectations indicate a more persistent dynamic quantile updating process, consistent with the lower variability and weaker identifying content of long-run inflation expectations. Third, the lower estimates of γ under the specifications with long-run expectations indicate a weaker response of the positive-valued processes around the reference quantile to new information, suggesting more muted distributional dynamics.

Distinct persistence dynamics also emerge in the TVP updating processes under the time-varying DMQ specifications. Specifically, the homogeneous slope dynamics become highly persistent, approaching a near random-walk process when

Table 3.7: Estimated parameters of the DMQ specifications with short-run (S) and long-run (L) inflation expectations over 1961Q1–2025Q4

	DMQ	**DMQ(S)**	**DMQ(L)**	**DMQ-TV(S)**	**DMQ-TV(L)**
			Dynamic quantile update		
α	0.531	0.459	0.526	0.471	0.572
β	1.000	0.883	0.941	0.925	0.940
γ	0.070	0.083	0.042	0.080	0.034
ϕ	0.998	0.997	0.998	0.997	0.998
			TVP-related parameters		
ρ_κ	–	–	–	0.013	0.281
ϕ_κ	–	–	–	0.756	0.998
ρ_β	–	–	–	0.088	0.302
ϕ_β	–	–	–	0.985	0.849
			Semi-structural parameters		
$1-\beta$	<0.001	0.117	0.059	0.075	0.060
κ	–	0.052	0.010	0.052	0.005
			Model fit		
Loglik.	6017.1	5886.2	5975.4	5871.3	5936.8
AIC	12042.2	11782.4	11960.9	11756.6	11887.6

Note: Under DMQ-TV, the reported β values correspond to sample averages. Loglik. denotes the final negative quasi-maximum likelihood (QML) objective function evaluated at the estimated parameter values.

long-run expectations are considered. In contrast, slope-updating persistence is less pronounced in the DMQ-TV(S) specification.

The opposite pattern emerges for the expectation-weight process, with stronger inertia in the relative weight assigned to short-run expectations compared with the specification based on long-run expectations.

A similar pattern arises in the dynamic quantile and TVP updating processes in the sDMQ specifications reported in Table 3.8. Notably, the average values of the slope parameters κ become more stable and more comparable across specifications. This suggests that removing the homogeneity constraint on the slope parameter alleviates potential identification issues associated with long-run inflation expectations. More importantly, the estimated slope parameters across the target probability levels reveal non-negligible heterogeneity, driven primarily by the substantially lower slope estimates in the upper tail of the conditional distribution.

Turning to the sDMQ-TV models, the high value of the shrinkage parameter ψ_s implies that the time- and quantile-averaged values of the time-varying slope parameter $\kappa_{p,t}$ remain comparable to those obtained under the static sDMQ specifications. Notably, the sDMQ-TV specification with short-run expectations attains the lowest aggregate loss and information criteria, indicating superior overall model fit.

Figure 3.5 provides a visual overview of the estimated slopes and reveals pronounced asymmetry. The slope is consistently steepest at the lower tail (long-run specifications) and the first quartile (short-run specifications) of the inflation distribution. This pattern suggests that inflation is more responsive to real activity under lower-tail and lower-quartile inflation outcomes. Moving toward the median, the slope flattens somewhat but remains moderately steep around the median, before declining sharply toward the upper tail, indicating a weak output-inflation

Table 3.8: Estimated parameters of the sDMQ specifications with short-run (S) and long-run (L) inflation expectations over 1961Q1–2025Q4

	sDMQ(S)	**sDMQ(L)**	**sDMQ-TV(S)**	**sDMQ-TV(L)**
		Dynamic quantile update		
α	0.509	0.564	0.483	0.568
β	0.872	0.939	0.922	0.937
γ	0.084	0.025	0.075	0.024
ϕ	0.997	0.998	0.998	0.998
		TVP-related parameters		
ρ_κ	–	–	0.112	0.148
ϕ_κ	–	–	0.712	0.920
ρ_β	–	–	0.100	0.335
ϕ_β	–	–	0.985	0.849
ψ_s	–	–	0.969	0.839
		Semi-structural parameters		
$1-\beta$	0.128	0.117	0.078	0.063
κ	0.069	0.052	0.066	0.057
$\kappa_{\tau=0.05}$	0.092	0.093	0.091	0.081
$\kappa_{\tau=0.25}$	0.094	0.088	0.093	0.073
$\kappa_{\tau=0.50}$	0.077	0.081	0.076	0.067
$\kappa_{\tau=0.75}$	0.065	0.064	0.065	0.059
$\kappa_{\tau=0.95}$	0.003	0.003	0.005	0.006
		Model fit		
Loglik.	5845.6	5939.4	5831.1	5914.3
AIC	11709.1	11896.8	11676.1	11842.7

Note: Under the sDMQ specification, the reported values of κ are averages across quantiles, while under sDMQ-TV they are averaged across both time and quantiles. Under sDMQ-TV, both κ_p and β correspond to sample averages, and ϕ_β is obtained from DMQ-TV. The reported Loglik. values correspond to the final value of negative quasi-maximum likelihood (QML) objective function evaluated at the estimated parameter values.

trade-off at higher inflation outcomes.

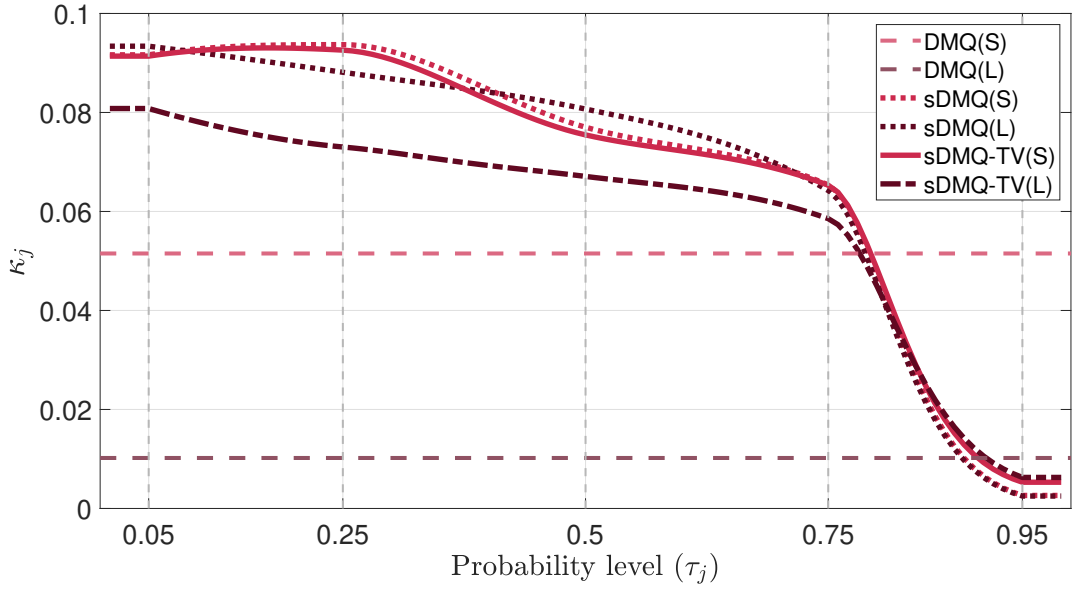

Figure 3.5: Phillips curve slopes across inflation quantiles obtained from different specifications. Values for the sDMQ-TV model correspond to sample-averaged slopes.

It should be noted that the static sDMQ model yields less stable estimates of the

slope parameter in the lower tail, as reflected in their sensitivity to initial values and estimation settings. Such instability likely also contributes to the distinct lower-tail slope estimates obtained in the benchmark analysis presented in Section 3.3. To address this, we conduct a robustness analysis in which the target tail quantiles are set to $\tau = 0.1$ and $\tau = 0.9$, with flat extrapolation applied beyond these thresholds to increase the effective number of observations in the tails. While the resulting estimates are qualitatively similar, we acknowledge the possibility of weak identification in the lower tail under the static sDMQ specification.

The ambiguity surrounding the static lower-tail slope parameter under the sDMQ specification likely reflects multiple factors, including the limited number of effective observations due to the rarity of deflationary episodes in the sample period and potential time variation. Given the improved model fit of the DMQ-TV and sDMQ-TV specifications relative to more restricted alternatives, the evidence points to time variation as an important contributing factor. Overall, these results indicate that the inflation–output gap trade-off differs systematically across the inflation distribution and evolves over time, highlighting the importance of modeling time-varying and quantile-heterogeneous slopes rather than relying on static homogeneous specifications.

In summary, the results indicate clear heterogeneity in the trade-off between inflation and the output gap across the inflation distribution. The static sDMQ model consistently reveals an asymmetric slope profile, with the steepest slopes at lower and central quantiles and a sharp decline toward the upper tail. In general, this pattern is preserved when allowing for time variation in the slope parameters. Model fit improves substantially under the sDMQ specifications relative to the benchmark DMQ models, with the sDMQ-TV specifications delivering the best overall fit.

3.4.6 Phillips curve slope dynamics

This section analyzes the evolution of the time-varying Phillips curve slope across quantiles of the conditional inflation distribution. By allowing the inflation–output gap trade-off to vary over time and across probability levels, we assess the extent of time variation and distributional asymmetry in the slope dynamics.

Figure 3.6 displays the time-varying Phillips curve slope estimates obtained from the sDMQ-TV specification with short-run inflation expectations, revealing important dynamics in the shape of the Phillips curve that remain hidden when considering its static distributional profile alone. The slope is steepest and exhibits the greatest variability at the first quartile and in the lower tail; it remains moderate and relatively stable around the median and declines toward higher probability levels, becoming nearly flat in the upper tail.

The results indicate an asymmetric Phillips curve across the conditional inflation distribution, with time-varying patterns that coincide with major macroeconomic episodes in the US economy. Importantly, this time variation alters the slope profile substantially in certain periods. For example, during the early 1980s, under Volcker monetary tightening, the slope asymmetry becomes more pronounced, with relatively steep slopes at lower and middle quantiles. As the monetary policy shock attenuates, however, the slope profile compresses, leading to more homogeneous slopes across quantiles, except for the upper tail.

Another notable change in the slope profile occurs in the post-GFC period, when the slope at the first quartile reaches historically high levels. This pronounced shift

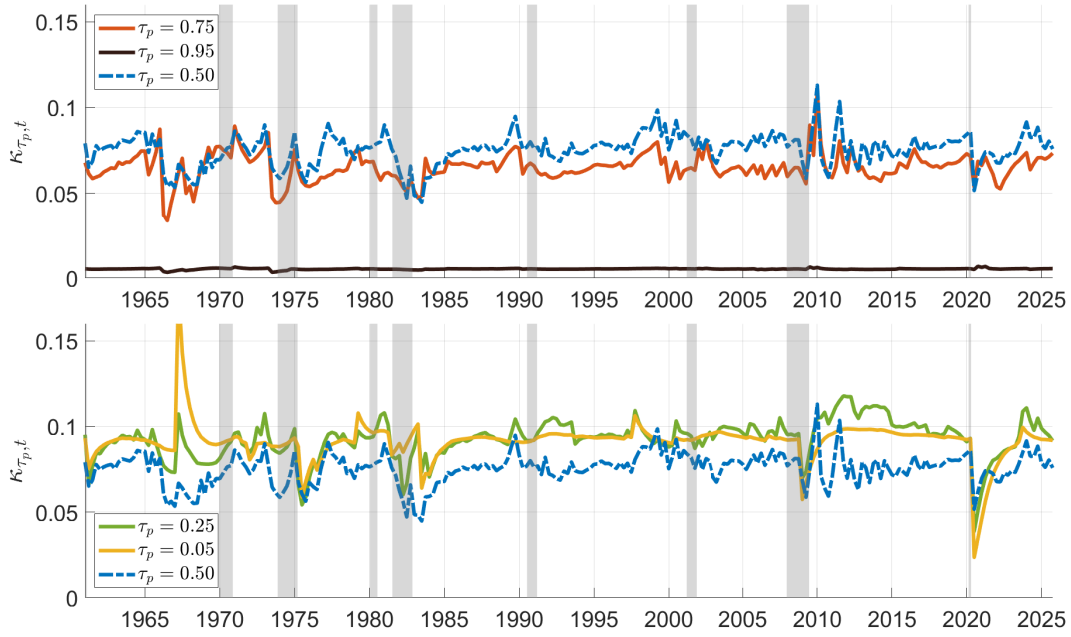

Figure 3.6: Time-varying Phillips curve slope parameters across the US conditional inflation distribution at the target probability levels obtained by sDMQ-TV specification with short-run expectations.

coincides with periods of effective lower bound and below-target inflation. From a New Keynesian perspective, this pattern is consistent with state-dependent Phillips curve dynamics, potentially reflecting greater inflation sensitivity to real marginal costs under low-inflation conditions and constrained monetary policy. Given the effective lower bound and the pronounced steepening at the first quartile, this pattern suggests that mechanisms beyond downward nominal rigidities were likely at work during this period, consistent with Fuhrer et al. (2012).

Lastly, following the COVID-19 pandemic, the Phillips curve relationship weakens markedly in the lower quantiles of the inflation distribution. This breakdown in the link between the output gap and inflation is consistent with the absence of disinflation. In contrast, standard New Keynesian models with a time-invariant Phillips curve slope would typically predict disinflation following a sharp but transitory contraction in output.

Figure 3.7 displays the estimated slopes based on the specification with long-run expectations. While the slope profile dynamics are broadly comparable to those observed under the short-run expectations specification, their time variation is more pronounced. Specifically, except for the post-Volcker disinflationary episode, the degree of heterogeneity exhibits greater variability and more substantial cross-quantile shifts over time. Importantly, the collapse of the Phillips curve relationship in the lower tail and at the first quartile during the post-COVID period becomes more evident relative to the specification with short-run expectations.

Overall, comparable Phillips curve dynamics emerge under both the long- and short-run expectations specifications, albeit with stronger persistence and greater variability under the long-run specification. The flattening of the Phillips curve in the lower tail and at the first quartile following the Great Recession and during the post-COVID period is consistent with the missing disinflation observed during these episodes. More broadly, the results highlight the importance of allowing for both time variation and distributional heterogeneity in the analysis of US inflation

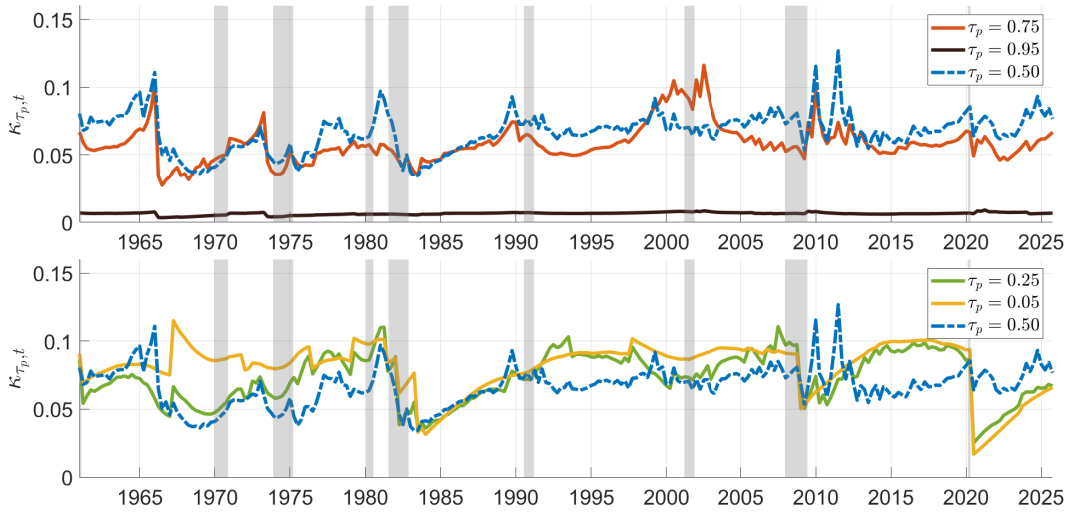

Figure 3.7: Time-varying Phillips curve slope parameters across the US conditional inflation distribution at the target probability levels obtained by sDMQ-TV specification with long-run expectations.

dynamics. The sDMQ-TV specification captures economically meaningful variation that is obscured in models imposing static or quantile-invariant slopes. The Phillips curve exhibits an asymmetric shape-profile across quantiles that evolves over time, with changes in both magnitude and shape across macroeconomic episodes. These findings suggest that conflicting empirical evidence on the Phillips curve flattening hypothesis may reflect substantial heterogeneity, asymmetry, and time variation across the inflation distribution.

3.4.7 Time-varying relative weight of inflation expectations

Figure 3.8 displays the time-varying weight on short-run and long-run inflation expectations, measured as $1 - \beta_t$, where β_t is the AR parameter governing inflation inertia. Higher values of $1 - \beta_t$ therefore indicate a stronger forward-looking component and a weaker role for backward-looking information in shaping inflation dynamics.

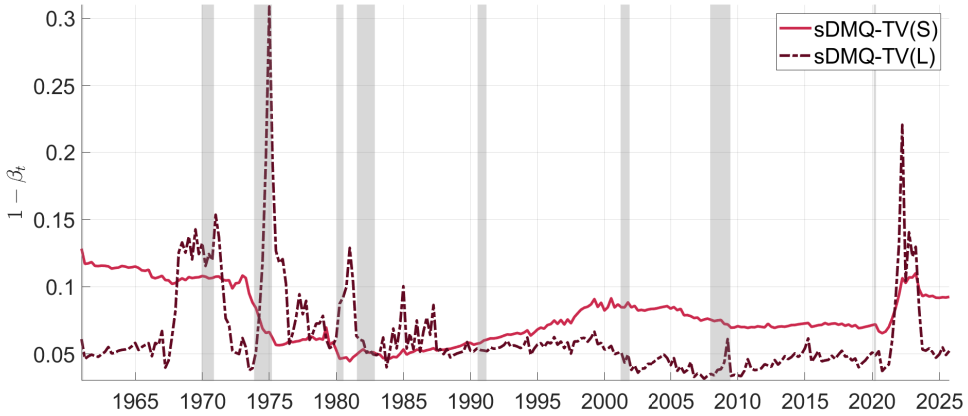

Figure 3.8: Time-varying relative weight on short-run and long-run forward-looking inflation expectation obtained from the sDMQ-TV specification.

The figure reveals regime dependence in the role of both short-run and long-run

inflation expectations. In the late 1960s and early 1970s, the weight on expectations is relatively elevated, particularly for long-run expectations, consistent with weakly anchored expectations that amplified inflation dynamics and contributed to the Great Inflation. The fact that these periods are not systematically associated with steeper Phillips curve slopes supports the view that expectations played a central role in sustaining high inflation outcomes.

Comparing the two measures, the specification based on short-run expectations displays more persistent dynamics, while the specification based on long-run expectations exhibits sharper and more episodic movements, especially during periods of macroeconomic stress. Specifically, the weight on long-run expectations shows pronounced spikes during the 1970s inflationary episode and again in the post-2020 period, whereas the weight on short-run expectations evolves more gradually and remains at comparatively moderate levels. In the specification based on short-run expectations, the observed strong persistence with moderate time variation is consistent with earlier evidence showing that persistence in PCE and GDP deflator inflation remained comparatively high and stable since the 1960s (Pivetta and Reis, 2007; Benati, 2008; Wolters and Tillmann, 2015). This pattern suggests that price-setting behavior relies more heavily on long-run expectations during inflationary episodes, while the relative weight on short-run expectations remains more stable over time.

3.4.8 Dynamics of the US inflation quantiles

This section reports the dynamic conditional inflation quantiles for the sDMQ-TV models with both short-run and long-run inflation expectations.

By tracking the evolution of selected target quantiles, we analyze time variation in the conditional distribution of US inflation. Focusing on distributional dynamics rather than mean outcomes allows us to identify shifts in asymmetry, dispersion, and tail behavior across macroeconomic regimes.

The filtered conditional target inflation quantiles obtained with short-run expectations are reported in Figure 3.9. A clear pattern emerges in the shape of the distribution when comparing the five target quantiles over time. These distributional shifts are consistent with well-documented episodes in US macroeconomic history.

The oil shocks of the 1970s pushed the right tail of the inflation distribution markedly away from the median, consistent with elevated inflationary risk. The right tail asymmetry remained substantial for an extended period, reflecting strong inflation persistence and de-anchoring inflation expectations combined with a relatively weak sensitivity to output gap fluctuations above the central quantiles of the inflation distribution.

After the monetary tightening of the 1980s, consistent with consolidating monetary policy during the Great Moderation, the inflation quantiles became more tightly clustered, reflecting lower volatility during this interval. However, this period of relative stability was disrupted by the GFC. The widening spread between quantiles around the median reflects the increase in uncertainty during this period.

More recently, a clear contrast emerges between the pre- and post-GFC decades. While realized PCE inflation after the GFC was only marginally lower than in the preceding period, the lower-tail quantiles at $\tau = 0.05$ periodically became negative. This indicates that, despite the Fed's ability to keep average inflation between 1% and its 2% PCE inflation target, deflationary risks remained non-negligible in the

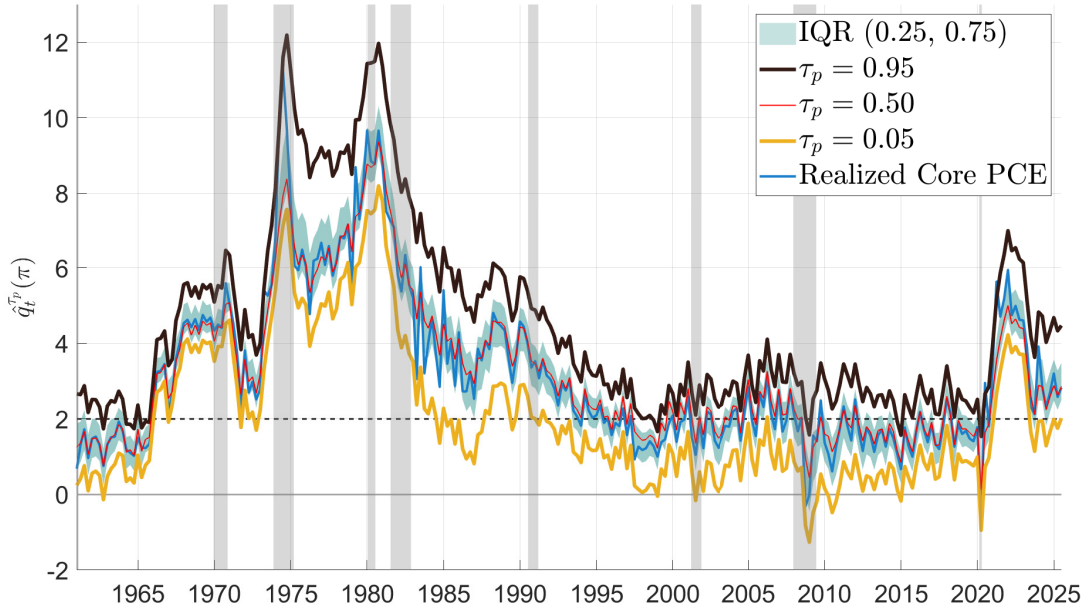

Figure 3.9: Filtered inflation quantiles obtained by the sDMQ-TV model with short-run expectations together with realized core CPI inflation. The horizontal dashed line represents the Fed's 2% PCE inflation target, shown for reference. Vertical shaded areas indicate recessions as identified by the National Bureau of Economic Research (NBER).

post-GFC environment, consistent with concerns raised by Fuhrer et al. (2012).

The severe macroeconomic shocks associated with the COVID-19 pandemic marked a turning point in inflation dynamics. After a brief period of downward pressure on prices, the subsequent demand surge ended a prolonged period characterized by relatively low inflationary risk. Importantly, the right-tail quantiles remained above 4%, indicating elevated inflation-at-risk at the end of the analyzed period. Moreover, the weak Phillips curve relationship in the upper tail, together with strong inflation persistence, helps explain the prolonged elevation of inflationary risk observed during this period.

The filtered inflation quantiles obtained under the long-run expectation specification reported in Figure 3.10 display broadly similar distributional dynamics, but with some notable differences relative to the short-run specification. Specifically, the long-run expectation specification generates more compressed dynamics during inflationary episodes. While the upper-tail inflation quantiles are generally less persistent, the around 4% inflation-at-risk at the end of the analyzed period remains robust when long-run expectations are considered.

Overall, the result demonstrates that the filtered conditional inflation quantiles exhibit strong persistence, and tend to diverge during periods of macroeconomic stress, especially in inflationary episodes. Importantly, the analysis of inflation quantiles demonstrates that the sDMQ model effectively captures time-varying asymmetry and heteroskedasticity in the conditional distribution of the US inflation. The results show that inflation dynamics are asymmetric and evolve over time, with higher conditional variability in the right tail, particularly during recessions and high-inflation periods.

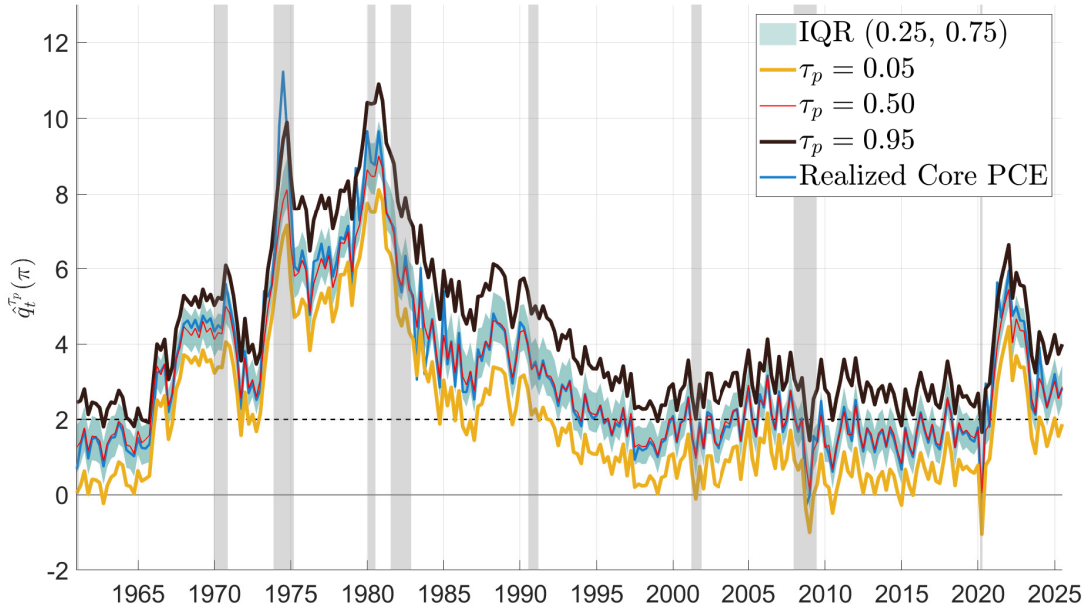

Figure 3.10: Filtered inflation quantiles obtained by the sDMQ-TV model with long-run expectations together with realized core CPI inflation. The horizontal dashed line represents the Fed's 2% PCE inflation target, shown for reference. Vertical shaded areas indicate recessions as identified by the National Bureau of Economic Research (NBER).

3.4.9 Time-varying conditional moments

The pronounced variability observed in the dynamic conditional inflation quantiles indicates that the US inflation distribution has undergone substantial changes over the past 65 years. However, direct quantitative interpretation of this variability based solely on the quantile dynamics is not straightforward.

To gain further insight, a skewed t-distribution, as proposed by Azzalini and Capitanio (2003), is fitted at each time t using the estimated target quantiles. The resulting implied conditional moments provide a convenient summary of the time-varying shape of the inflation distribution. Moreover, comparing the higher-order conditional moments obtained under specifications with short-run and long-run inflation expectations allows for a deeper assessment of the distributional implications of the differences in their dynamics.

The skew t density takes the form of

$$f(y;\mu,\sigma,\zeta,\nu) = \frac{2}{\sigma}\, d_T\left(\frac{y-\mu}{\sigma};\nu\right)\, T\left(\zeta\,\frac{y-\mu}{\sigma}\sqrt{\frac{\nu+1}{\nu+\left(\frac{y-\mu}{\sigma}\right)^2}}\,;\,\nu+1\right), \qquad (3.31)$$

where $d_T(\cdot)$ and $T(\cdot)$ denote the probability density function (PDF) and CDF of the skew t distribution, respectively. The parameters are location (μ), scale (σ), degrees of freedom (ν) governing tail thickness, and shape (ζ) governing the skewness. Estimating these parameters at each time t supports the assessment of the tail behavior of inflation by capturing its time-varying asymmetry and kurtosis implied by the estimated target quantiles.

Figure 3.11 displays the filtered time-varying conditional moments. The upper left panel shows the filtered conditional mean of inflation, which captures the low-frequency trend in core PCE by removing transitory shocks and extreme realizations. Realized core PCE exhibits pronounced short-run deviations around this mean, particularly during the 1970s inflation episode and the post-2020 inflation

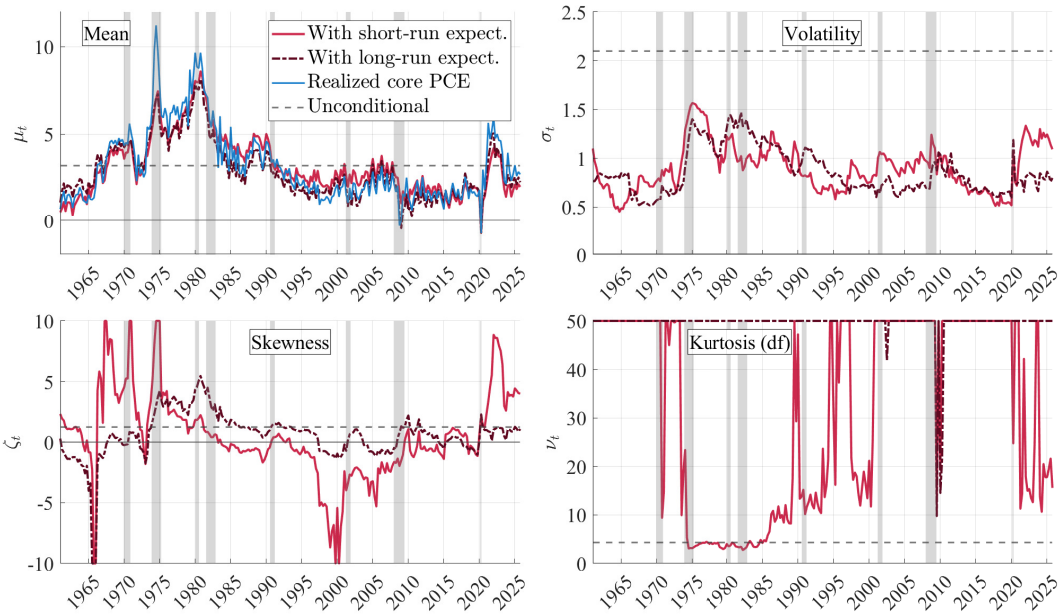

Figure 3.11: Filtered implied conditional moments obtained by fitting a skewed-t distribution, following Azzalini and Capitanio (2003), at each time t using the target quantiles. The parameters are the location (μ_t, upper left panel), the scale (σ_t, upper right panel), the shape parameter (ζ_t, lower left panel) governing skewness, and the degrees of freedom (ν_t, lower right panel) governing tail thickness. The horizontal dashed lines denote the corresponding unconditional moments. Shaded vertical areas indicate recessions as identified by the National Bureau of Economic Research (NBER).

surge. The specification with short-run expectations captures more cyclical variation, whereas the long-run specification yields a smoother and more persistent path.

The upper right panel reports the filtered conditional inflation volatility, which peaks during the 1970s and early 1980s and declines markedly thereafter. Volatility remains elevated during the high-inflation period and persistently low during the Great Moderation, with only a modest increase following the GFC and in the post-2020 period. The difference between the two specifications becomes more pronounced: volatility is more time-varying and shock-sensitive under short-run expectations, but smoother and more persistent under long-run expectations due to anchoring.

A similar pattern emerges for the conditional skewness reported in the lower left panel. Under short-run expectations, skewness is strongly time-varying, with large positive spikes during the 1970s and occasional negative phases in the late 1960s and the late 1990s. In contrast, under long-run expectations the conditional inflation distribution remains close to symmetric from the 1980s onward. This pattern is consistent with the Great Moderation and reflects a transition associated with monetary policy consolidation, better-anchored inflation expectations, and reduced inflationary risks.

Lastly, the lower right panel reports filtered kurtosis across specifications. Under long-run expectations, excess kurtosis largely vanishes, whereas it remains pronounced under short-run expectations. This indicates that the specification with more anchored expectations yields a distribution closer to Gaussian, while short-run expectations generate heavier tails.

In summary, the results highlight substantial differences in the implied distributional dynamics across expectation proxies. The specification with short-run

expectations yields richer time variation across all conditional moments, with more pronounced volatility, skewness, and excess kurtosis. This stronger time variation reflects the transmission of greater short-run expectation variability into the conditional distribution. In contrast, the specification with long-run expectations produces a smoother and more stable distribution, characterized by attenuated higher-order moments and dynamics closer to a Gaussian benchmark. Overall, these findings underscore the important role of expectations in shaping the dynamics and higher-order properties of the conditional distribution of inflation.

3.5 Conclusions

This chapter analyzed the US inflation dynamics by jointly modeling multiple quantiles within a hybrid New Keynesian Phillips curve framework. To this end, the Smoothed Dynamic Multiple Quantile model was introduced, allowing for quantile-specific and time-varying parameters. By linking the time variation of a reduced set of quantiles to the full conditional distribution within a score-driven framework, the proposed specification captures heterogeneous responses of inflation to real activity across the distribution and over time.

The empirical results revealed pronounced time-varying and distributional nonlinearities in the inflation–output gap trade-off, indicating clear heterogeneity across the conditional inflation quantiles. Specifically, the Phillips curve exhibits an asymmetric profile across quantiles, with relatively steep slopes at lower probability levels and a sharp flattening towards the upper tail. Allowing for time variation showed that this distributional pattern evolves over time, consistent with state-dependent New Keynesian price setting behavior.

More broadly, the proposed framework provides a flexible and structurally interpretable tool for analyzing inflation dynamics across different parts of the distribution and over time. The results highlight the central role of inflation expectations in shaping high-inflation episodes, particularly during the Great Inflation, while emphasizing the dominance of persistence and weak real activity pass-through in the lower quantiles following the Global Financial Crisis and during the COVID-19 pandemic. The combination of a flattening Phillips curve slope in the lower quantiles and strong persistence helps explain the absence of disinflation during sharp output contractions, while persistently flat slopes in the upper tail explain the challenge of stabilizing inflation when inflationary risks are elevated.

In addition, the results underscore that the choice of expectation proxy plays a critical role for identification, with short-run expectations revealing richer distributional dynamics and more pronounced higher-order moments, while more anchored long-run expectations yield smoother distributions but stronger time variation in the cross-quantile Phillips curve shape.

Lastly, the filtered inflation quantiles provide a quantitative measure of inflationary risk, capturing both its persistence and asymmetric evolution over time. Notably, the upper tail conditional inflation quantile has remained around 4% since the COVID-19 pandemic indicating that inflationary risk remains a non-negligible factor at the end of the analyzed period. Importantly, the weak Phillips curve relationship in the upper tail, together with strong inflation persistence, helps explain the prolonged elevation of inflationary risk observed during this period.

Overall, the results suggest that exploiting distributional information and time variation can provide additional insights into inflation dynamics beyond those available from conventional limited-information empirical approaches.

More generally, the framework is not limited to inflation, and its generality allows it to be applied in other macro-financial contexts where time-varying quantile-heterogeneous parameters, persistence, and tail risks are central. In addition, its ability to model time-varying higher-order moments enables full distributional forecasting, providing a more comprehensive assessment of tail risks.

Summary

This thesis revolved around the development and application of score-driven models for analyzing volatility dynamics in financial markets and structural relationships among macroeconomic variables.

First focusing on financial applications, Chapter 1 introduced the Realized Dynamic Score EGARCH model with heavy-tailed, skewed conditional distributions for the joint modeling of returns and a realized range-based volatility measure. The proposed model employs a score-driven framework for jointly modeling asset returns and a range-based realized volatility measure. Additionally, the model incorporates skewed and heavy-tailed distributions, which are inherent features of financial time series but are often not well captured by traditional GARCH models. The framework also includes a two-component volatility structure and extended leverage channels, which increase flexibility and improve model fit. By exploiting the information contained in the conditional densities through their scores, together with additional enhancements, the model enables more responsive and adaptive modeling of financial volatility.

Empirical analysis based on in-sample fit and out-of-sample VaR and ES forecasts for the S&P 500 and the two considered individual assets indicates that the proposed models outperform beta-skew-t-EGARCH variants. These findings reflect improved model fit and a stronger ability to capture volatility dynamics. While not uniformly dominant, the MCS results provide consistent support for the proposed specification, particularly for the more parsimonious variants, which exhibit stable performance across loss functions and tail probabilities.

Given the longer data availability for the two individual assets, future research could investigate model performance over extended forecasting horizons and across broader market environments. In addition, extending the set of considered assets, particularly to those where asymmetric volatility responses and model flexibility are especially relevant, would provide a more comprehensive assessment of the framework.

Finally, the two-component structure of the model, which captures long-run dynamics, naturally motivates the evaluation of multi-step-ahead forecasts. Assessing such horizons would offer additional insight into the ability of the model to capture persistent volatility components and long-memory features.

By shifting the focus to macroeconomics, Chapter 2 investigated the dynamics of the natural rate of interest (r-star) in the US using a score-driven state-space model within the Laubach–Williams semi-structural framework. The main contribution of the chapter is to allow the parameters of the Laubach–Williams model to vary over time, with their dynamics governed by the conditional likelihood score. In addition, motivated by the econometric challenges initiated by the pandemic, the chapter introduced an augmented accelerating version of the GAS (aaGAS) updating mechanism.

The output gap estimated by the aaGAS suggests that the modified model ef-

fectively handles large volatility shifts due to extreme events. Furthermore, the higher estimate of potential output relative to HLW suggests that the proposed models are more efficient at disentangling transitory and highly persistent shocks. Regarding the time-varying structural relationships, the IS slope was highly unstable until the beginning of the Great Moderation, gradually flattening during this period and stabilizing at a low level after the Great Recession. As for the Phillips curve, while we find little evidence of time variation in its slope parameter, the pronounced downside uncertainty in the simulated parameter limits any decisive conclusions regarding the stability of the slope over the past decades. Most importantly, the r-star estimates exhibit substantially higher precision and reduced variability, along with a flatter downward trend over the analyzed interval compared to the benchmark HLW model. In addition, we find that the primary source of the recent divergence in the trajectories and the uncertainty surrounding the natural rate of interest lies in the estimation of the "other factors", z. These state estimates not only deviate substantially from those of HLW but also account for most of the improvement in r-star estimation precision, explaining the higher natural rate of interest—ranging between 1.5% and 2%—observed over the past two decades. Furthermore, the current upward trend in both the filtered and smoothed estimates, in contrast to HLW model output, indicates that the natural rate of interest has already passed its historically low level.

Given the high uncertainty and the ambiguous definition of the z-factor, future research could focus on improving its refined identification and developing more precise specifications. Moreover, considering the low-frequency nature of current estimations and the delayed availability of relevant data, a model that utilizes higher-frequency data and proxy variables could offer a more timely and accurate assessment.

The lower variability of the r-star estimate, along with the reduced standard errors of the underlying state estimates, implies that the proposed models are more adaptable to sudden changes in the economic environment and, thus, more effective at identifying transitory shocks than the standard LW methodology. Based on a purely statistical, data-driven approach, the GAS (score-driven) models, therefore, provide a valid empirical alternative for estimating the natural rate of interest in real-time, demonstrating a clear advantage in estimation precision. In addition, the aaGAS specification further improves the precision of the r-star estimate, alleviating a long-standing limitation of models based on the LW methodology. Moreover, in light of the occurrence of extreme events in both macroeconomic and financial contexts, the aaGAS offers a viable extension of the standard GAS framework to more adequately model volatility persistence.

Motivated by the pronounced estimation uncertainty surrounding the Phillips curve slope parameter reported in Chapter 2, Chapter 3 analyzed the conditional distribution of US inflation by jointly modeling multiple inflation quantiles within a score-driven framework based on the New Keynesian Phillips Curve. To this end, the Smoothed Dynamic Multiple Quantile model was introduced, allowing for quantile-specific and time-varying parameters. By linking the time variation of a reduced set of quantiles to the full conditional distribution within a score-driven framework, the proposed specification captures heterogeneous responses of inflation to real activity across the distribution and over time.

The empirical results revealed pronounced time-varying and distributional nonlinearities in the inflation–output gap trade-off, indicating clear heterogeneity across the conditional inflation quantiles. Specifically, the Phillips curve exhibits an asym-

metric profile across quantiles, with relatively steep slopes at lower probability levels and a sharp flattening towards the upper tail. Allowing for time variation showed that this distributional pattern evolves over time, consistent with state-dependent New Keynesian price setting behavior.

More broadly, the proposed framework provides a flexible and structurally interpretable tool for analyzing inflation dynamics across different parts of the distribution and over time. The results highlight the central role of inflation expectations in shaping high-inflation episodes, particularly during the Great Inflation, while emphasizing the dominance of persistence and weak real activity pass-through in the lower quantiles following the Global Financial Crisis and during the COVID-19 pandemic. The combination of a flattening Phillips curve slope in the lower quantiles and strong persistence helps explain the absence of disinflation during sharp output contractions, while persistently flat slopes in the upper tail explain the challenge of stabilizing inflation when inflationary risks are elevated.

In addition, the results underscore that the choice of expectation proxy plays a critical role for identification, with short-run expectations revealing richer distributional dynamics and more pronounced higher-order moments, while more anchored long-run expectations yield smoother distributions but stronger time variation in the cross-quantile Phillips curve shape.

Lastly, the filtered inflation quantiles provide a quantitative measure of inflationary risk, capturing both its persistence and asymmetric evolution over time. Notably, the upper tail conditional inflation quantile has remained around 4% since the COVID-19 pandemic indicating that inflationary risk remains a non-negligible factor at the end of the analyzed period. Importantly, the weak Phillips curve relationship in the upper tail, together with strong inflation persistence, helps explain the prolonged elevation of inflationary risk observed during this period.

Overall, the results suggest that exploiting distributional information and time variation can provide additional insights into inflation dynamics beyond those available from conventional limited-information empirical approaches.

More generally, the framework is not limited to inflation, and its generality allows it to be applied in other macro-financial contexts where time-varying quantile-heterogeneous parameters, persistence, and tail risks are central. In addition, its ability to model time-varying higher-order moments enables full distributional forecasting, providing a more comprehensive assessment of tail risks.

In conclusion, applications of the developed score-driven models to financial and macroeconomic problems demonstrated that, with appropriate refinements, the score-driven framework offers a coherent and flexible approach for capturing nonlinear dynamics inherent in complex data generating processes.

Bibliography

Adrian, T., N. Boyarchenko, and D. Giannone (2019, April). Vulnerable growth. *American Economic Review 109*(4), 1263–89.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*(2), 579–625.

Azzalini, A. and A. Capitanio (2003, May). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society Series B 65*(2), 367–389.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 74*(1), 3–30.

Ball, L. and S. Mazumder (2011). Inflation Dynamics and the Great Recession. *Brookings Papers on Economic Activity 42*(1 Spring), 337–405.

Bauer, M. D., B. S. Bernanke, and E. Milstein (2023, February). Risk appetite and the risk-taking channel of monetary policy. *Journal of Economic Perspectives 37*(1), 77–100.

Bauwens, L. and N. Hautsch (2006). Stochastic conditional intensity processes. *Journal of Financial Econometrics 4*(3), 450–493.

Benati, L. (2008). Investigating inflation persistence across monetary regimes. Technical Report Working Paper No. 851, European Central Bank.

Benigno, P. and G. B. Eggertsson (2023). It's baaack: The surge in inflation in the 2020s and the return of the non-linear phillips curve. Working Paper 31197, National Bureau of Economic Research.

Bernanke, B. and M. Gertler (1989, March). Agency costs, net worth, and business fluctuations. *American Economic Review 79*(1), 14–31.

Bernanke, B. S., M. Gertler, and S. Gilchrist (1999). Chapter 21 the financial accelerator in a quantitative business cycle framework. Volume 1 of *Handbook of Macroeconomics*, pp. 1341–1393. Elsevier.

Blasques, F., C. Francq, et al. (2023). Quasi score-driven models. *Journal of Econometrics 231*(2), 235–257.

Blasques, F., P. Gorgi, and S. Koopman (2019). Accelerating score-driven time series models. *Journal of Econometrics 212*(2), 359–376.

Blasques, F., S. J. Koopman, K. Łasak, and A. Lucas (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.

Bloom, N., P. Bunn, P. Mizen, P. Smietanka, and G. Thwaites (2025). The impact of COVID-19 on productivity. *The Review of Economics and Statistics 107*(1), 28–41.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics 69*(3), 542–547.

Borio, C. (2014). The financial cycle and macroeconomics: What have we learnt? *Journal of Banking & Finance 45*, 182–198.

Buncic, D. (2024). Econometric issues in the estimation of the natural rate of interest. *Economic Modelling 132*, 106641.

Caporin, M. (2008, 12). Evaluating value-at-risk measures in presence of long memory conditional volatility. *Journal of Risk 10*.

Catania, L. and A. Luati (2023). Semiparametric modeling of multiple quantiles. *Journal of Econometrics 237*(2, Part B), 105365.

Christensen, K. and M. Podolskij (2007). Realized range-based estimation of integrated variance. *Journal of Econometrics 141*(2), 323–349.

Christensen, K., M. Podolskij, and M. Vetter (2009, April). Bias-correcting the realized range-based variance in the presence of market microstructure noise. *Finance and Stochastics 13*(2), 239–268.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review 39*(4), 841–862.

Clark, T. E. and T. Doh (2014). Evaluating alternative models of trend inflation. *International Journal of Forecasting 30*(3), 426–448.

Cochrane, J. H. (2011). Determinacy and identification with taylor rules. *Journal of Political Economy 119*(3), 565–615.

Cox, D. R. (1981). Statistical analysis of time series: Some recent developments. *Scandinavian Journal of Statistics 8*(2), 93–115.

Creal, D., S. J. Koopman, and A. Lucas (2008). A general framework for observation driven time-varying parameter models. Technical Report 08-108/4, Tinbergen Institute.

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

Curdia, V. and M. Woodford (2010). Credit spreads and monetary policy. *Journal of Money, Credit and Banking 42*(S1), 3–35.

Curdia, V. and M. Woodford (2011). The central-bank balance sheet as an instrument of monetary policy. *Journal of Monetary Economics 58*(1), 54–79.

Curdia, V. and M. Woodford (2016). Credit frictions and optimal monetary policy. *Journal of Monetary Economics 84*, 30–65.

Davis, R. A., W. T. M. Dunsmuir, and S. B. Streett (2003). Observation-driven models for poisson counts. *Biometrika 90*(4), 777–790.

Del Negro, M., D. Giannone, M. P. Giannoni, and A. Tambalotti (2017). Safety, liquidity, and the natural rate of interest. *Brookings Papers on Economic Activity*, 235–294.

Delle Monache, D., I. Petrella, and F. Venditti (2021). Price dividend ratio and long-run stock returns: A score-driven state space model. *Journal of Business & Economic Statistics 39*(4), 1054–1065.

Ding, Z. and C. W. Granger (1996). Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics 73*(1), 185–215.

Dupraz, S., E. Nakamura, and J. Steinsson (2019). A Plucking Model of Business Cycles. NBER Working Papers 26351, National Bureau of Economic Research, Inc.

Eggertsson, G. B. and P. Krugman (2012, None). Debt, deleveraging, and the liquidity trap: A fisher-minsky-koo approach. *The Quarterly Journal of Economics 127*(3), 1469–1513.

Engle, R. (2002). Dynamic conditional correlation. *Journal of Business & Economic Statistics 20*(3), 339–350.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica 50*(4), 987–1007.

Engle, R. F. and G. G. J. Lee (1999). A permanent and transitory component model of stock return volatility. *Handbook of Statistics 14*, 475–497.

Engle, R. F. and S. Manganelli (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics 22*(4), 367–381.

Engle, R. F. and J. R. Russell (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica 66*(5), 1127–1162.

Fernald, J. and H. Li (2022). The impact of COVID on productivity and potential output. *Federal Reserve Bank of San Francisco, Working Paper Series 2022*, 01–52.

Fernández, C. and M. F. J. Steel (1998). On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association 93*(441), 359–371.

Fisher, I. (1932). *Booms and depressions: some first principles*. Adelphi New York.

Friedman, M. (1993). The "plucking model" of business fluctuations revisited. *Economic Inquiry 31*(2), 171–177.

Fritsch, F. N. and R. E. Carlson (1980). Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis 17*(2), 238–246.

Fuhrer, J. C., G. P. Olivei, and G. M. B. Tootell (2012). Inflation dynamics when inflation is near zero. *Journal of Money, Credit and Banking 44*(s1), 83–122.

Furceri, D., D. Giannone, M. F. Kisat, M. W. R. Lam, and H. Li (2025, May). Debt-at-risk. IMF Working Papers 2025/086, International Monetary Fund.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica 79*(3), 733–772.

Garnier, J. and B.-R. Wilhelmsen (2009). The natural rate of interest and the financial cycle. *Journal of Economic Behavior & Organization 36*.

Gertler, M. and P. Karadi (2011). A model of unconventional monetary policy. *Journal of Monetary Economics 58*(1), 17–34.

Gertler, M. and N. Kiyotaki (2010). Financial intermediation and credit policy in business cycle analysis. In B. M. Friedman and M. Woodford (Eds.), *Handbook of Monetary Economics*, Volume 3, pp. 547–599. Elsevier.

Giot, P. and S. Laurent (2003). Value-at-risk for long and short trading positions. *Journal of Applied Econometrics 18*(6), 641–663.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance 48*(5), 1779–1801.

Gordon, R. (1998). Foundations of the goldilocks economy: Supply shocks and the time-varying nairu. *Brookings Papers on Economic Activity 29*(2), 297–346.

Hamilton, J. D. (1986). A standard error for the estimated state vector of a state-space model. *Journal of Econometrics 33*(3), 387–397.

Hansen, B. E. (1994). Autoregressive conditional density estimation. *International Economic Review 35*(3), 705–730.

Hansen, P. R., Z. Huang, and H. H. Shek (2012). Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics 27*(6), 877–906.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*(2), 453–497.

Harvey, A. and D. Palumbo (2023). Score-driven models for realized volatility. *Journal of Econometrics 237*(2, Part B), 105448.

Harvey, A. C. (1990). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Harvey, A. C. (2013). *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Econometric Society Monographs. Cambridge University Press.

Harvey, A. C. and S. J. Koopman (1992). Diagnostic checking of unobserved-components time series models. *Journal of Business & Economic Statistics 10*(4), 377–389.

Harvey, A. C. and G. Sucarrat (2014). EGARCH models with fat tails, skewness and leverage. *Computational Statistics and Data Analysis 76*, 320–338.

Hazell, J., J. Herreño, E. Nakamura, and J. Steinsson (2022). The slope of the Phillips curve: Evidence from u.s. states. *The Quarterly Journal of Economics 137*(3), 1299–1344.

Holston, K., T. Laubach, and J. C. Williams (2017). Measuring the natural rate of interest: International trends and determinants. *Journal of International Economics 108*, S59–S75.

Holston, K., T. Laubach, and J. C. Williams (2023, June). Measuring the natural rate of interest after covid-19. Staff Reports 1063, Federal Reserve Bank of New York.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering 82*(Series D), 35–45.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: Likelihood inference and comparison with arch models. *The Review of Economic Studies 65*(3), 361–393.

Kindleberger, C. P. (2000). *Manias, Panics, and Crashes: A History of Financial Crises* (4 ed.). Cambridge: Cambridge University Press.

Kleiber, C. and S. Kotz (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*. Wiley Series in Probability and Statistics. Hoboken, NJ: Wiley.

Koenker, R. and F. Machado, José A. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association 94*(448), 1296–1310.

Komunjer, I. (2005). Quasi-maximum likelihood estimation for conditional quantiles. *Journal of Econometrics 128*(1), 137–164.

Koopman, S., R. Bannink, and J. Penzer (1998). *Messy time series*, pp. 103–143. United States: JAI Press.

Kotz, S., J. Kozubowski, Tomasz and K. Podgórski (2001). *The Laplace Distribution and Generalizations*. Boston: Birkhäuser.

Krustev, G. (2019). The natural rate of interest and the financial cycle. *Journal of Economic Behavior & Organization 162*, 193–210.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives 3*(2), 73–84.

Laubach, T. (2001). Measuring the NAIRU: Evidence from seven economies. *The Review of Economics and Statistics 83*(2), 218–231.

Laubach, T. and J. C. Williams (2003). Measuring the natural rate of interest. *The Review of Economics and Statistics 85*(4), 1063–1070.

López-Salido, D. and F. Loria (2024). Inflation at risk. *Journal of Monetary Economics 145*, 103570.

Martens, M. and D. van Dijk (2007). Measuring volatility with the realized range. *Journal of Econometrics 138*(1), 181–207.

Mavroeidis, S., M. Plagborg-Møller, and J. H. Stock (2014, March). Empirical evidence on inflation expectations in the new keynesian phillips curve. *Journal of Economic Literature 52*(1), 124–88.

Minsky, H. P. (2015). *Can it happen again?: Essays on instability and finance*. Routledge.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica 59*(2), 347–370.

Opschoor, A., M. van der Wel, and D. van Dijk (2018). Combining high-frequency data and score-driven models for volatility forecasting. *Journal of Econometrics 202*(1), 118–140.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business 53*(1), 61–65.

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review 47*(2), 527–556.

Patton, A. J., J. F. Ziegel, and R. Chen (2019). Dynamic semiparametric models for expected shortfall (and value-at-risk). *Journal of Econometrics 211*(2), 388–413.

Phillips, A. W. (1958). The relation between unemployment and the rate of change of money wage rates in the united kingdom, 1861–1957. *Economica 25*(100), 283–299.

Pivetta, F. and R. Reis (2007). The persistence of inflation in the United States. *Journal of Economic Dynamics & Control 31*(4), 1326–1358.

Poiraud-Casanova, S. and C. Thomas-Agnan (2000). About monotone regression quantiles. *Statistics & Probability Letters 48*(1), 101–104.

Rossi, B., A. Inoue, and Y. Wang (2024, Jun). Has the phillips curve flattened? French Stata Users' Group Meetings 2024 22, Stata Users Group.

Staiger, D., J. H. Stock, and M. W. Watson (1996). How precise are estimates of the natural rate of unemployment? Working Paper 5477, National Bureau of Economic Research.

Stock, J. H. (1994). Unit roots, structural breaks and trends. Volume 4 of *Handbook of Econometrics*, pp. 2739–2841. Elsevier.

Stock, J. H. and M. W. Watson (1998). Median unbiased estimation of coefficient variance in a time-varying parameter model. *Journal of the American Statistical Association 93*(441), 349–358.

Storn, R. and K. Price (1997). Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization 11*(4), 341–359.

Suah, J. L. (2024). The "plucking" model of the unemployment rate floor: Cross-country estimates and empirics. BIS Working Papers 1159, Bank for International Settlements.

Taylor, J. B. (1993). Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy 39*, 195–214.

Taylor, S. J. (1994). Modeling stochastic volatility: A review and comparative study. *Mathematical Finance 4*(2), 183–204.

White, H., T.-H. Kim, and S. Manganelli (2015). Var for var: Measuring tail dependence using multivariate regression quantiles. *Journal of Econometrics 187*(1), 169–188.

Wicksell, K. (1936). *Interest and Prices: A Study of the Causes Regulating the Value of Money*. Macmillan. Originally published in German as *Geldzins und Güterpreise* in 1898.

Wolters, M. H. and P. Tillmann (2015). The changing dynamics of US inflation persistence: a quantile regression approach. *Studies in Nonlinear Dynamics & Econometrics 19*(2), 161–182.