Dipartimento Economia e Statistica Cognetti de Martiis

Working Paper Series

12/22

FORECASTING INATION: A GARCH-IN-MEAN-LEVEL MODEL WITH TIME VARYING PREDICTABILITY

ALESSANDRA CANEPA, MENELAOS KARANASOS, ATHANASIOS PARASKEVOPOULOS and EMILIO ZANETTI CHINI

Department of Economics and Statistics "Cognetti de Martiis"
Campus Luigi Einaudi, Lungo Dora Siena 100/A, 10153 Torino (Italy)
www.est.unito.it

UNIVERSITÀ DEGLI STUDI DI TORINO

The Department of Economics and Statistics "Cognetti de Martiis" publishes research papers authored by members and guests of the Department and of its research centers. ISSN: 2039-4004

Forecasting Inflation: A GARCH-in-Mean-Level Model with Time Varying Predictability

A. Canepa*
University of Turin
and Brunel University London

M. G. Karanasos
Brunel University London

A. G. Paraskevopoulos
University of Piraeus

Emilio Zanetti Chini
University of Bergamo

September 16, 2022

Abstract

In this paper we employ an autoregressive GARCH-in-mean-level process with variable coefficients to forecast inflation and investigate the behavior of its persistence in the United States. We propose new measures of time varying persistence, which not only distinguish between changes in the dynamics of inflation and its volatility, but are also allow for feedback between the two variables. Since it is clear from our analysis that predictability is closely interlinked with (first-order) persistence we coin the term *persistapredictability*. Our empirical results suggest that the proposed model has good forecasting properties.

Keywords: GARCH-in Mean, Inflation persistence, Optimal forecasts, Structural breaks.

JEL Classification: C13, C22, C32, E17, E31, E5

*Address for correspondence: Alessandra Canepa, Department of Economics and Statistics Cognetti De Martiis, University of Turin, 10153 Via Lungo Dora Siena 100A, Turin, Italy. Email: Alessandra.Canepa@unito.it.

1 Introduction

The behavior of inflation has long been an object of interest to economists, especially to central banks, which are bounded by statutory mandate to maintain price stability. In the literature this debate can be streamlined in two strands of empirical research, each drawing on different background. Namely the problem of modelling inflation persistence and the impact of inflation uncertainty on inflation levels.

Broadly speaking, inflation persistence measures the speed at which the inflation rate returns to its equilibrium level after an inflationary shock: the faster inflation returns to its equilibrium level after a macroeconomic shock, the more effective monetary policy action can be, all else equal. As a result, optimal monetary policy crucially depends on the knowledge of inflation dynamics. For example, high inflation persistence may require a bolder monetary policy action to bring inflation under control. On the other side, a low level of inflation persistence may require a weaker or no action by monetary authorities in response to an exogenous shock. No wonder why the issue of modelling inflation persistence has been intensively investigated in empirical studies. In the literature, a large number of works make use of autoregressive (AR) model-based measures such as the largest autoregressive root (LAR) and the sum of the autoregressive coefficients (SAR) to measure persistence. However, empirical studies are often in considerable disagreement regarding the characteristic features of inflation dynamics. For instance, in his seminal paper Taylor (2000) found that in the United States the inflation persistence during the Volcker-Greenspan era was substantially lower than during the previous two decades. Similarly, Levin and Piger (2004) showed that high inflation persistence was not an inherent characteristic of industrial economies over the period 1984-2002. On the other hand, the work based on the SAR approach by Batini (2006) suggested relatively little evidence of shifts in inflation persistence for the Euro area (see also Pivetta and Reis, 2007; Stock, 2001, O'Reilly and Whelan, 2005; and Gerlach and Tillman, 2012).

Research economists have also placed a lot of emphasis on the relationship between inflation and its uncertainty.[1] In his seminal paper Friedman (1977) argued that nominal uncertainty causes an adverse output effect. Friedman's famous argument about the negative welfare effects of inflation consists of two claims: higher inflation increases nominal uncertainty, which then decreases output growth. Ball (1990) took Fridman's point further by developing a repeated game model in which (through the presence of two types of policymakers with different preferences, who stochastically alternate in power) higher inflation generates higher nominal uncertainty. Causality in the opposite direction, namely from inflation uncertainty to inflation, is instead suggested by the model proposed by Cukierman and Meltzer (1986), in which there is an incentive for policymakers to create inflation surprises to raise output growth. In

[1] Baker et al. (2016) offer a review of the literature on economic uncertainty on the overall macroeconomy. Interesting works on the impact of uncertainty on the economic system include Bloom (2009), Bachman *et al.* (2013), Bloom et al. (2014), and Scotti (2016). A closely related literature focuses explicitly on policy uncertainty. Early studies on the detrimental economic effects of monetary, fiscal, and regulatory policy uncertainty include the work by the Rodrik (1991), Higgs (1997), and Hassett and Metcalf (1999). More recently, Born and Pfeifer (2014) and Fernandez-Villaverde at al. (2015) study policy uncertainty in DSGE models, finding moderately negative effects, while Pastor and Veronesi (2012) model the theoretical links among fluctuations, policy uncertainty, and stock market volatility.

particular, according to Cukierman and Meltzer (1986) in the presence of uncertainty about the rate of monetary growth and, therefore, inflation, policymakers apply an expansionary monetary policy in order to surprise the agents and enjoy output gains. The argument that Central Banks tend to create inflation surprises in the presence of more inflation uncertainty implies a positive causal effect from nominal uncertainty to inflation.[2] Empirical research has sought evidence of the causal relations between inflation and inflation uncertainty by estimating GARCH-type models. One of the first papers to test for the Cukierman and Meltzer hypothesis in a context of a GARCH in-mean model was Baillie *et al.* (1996). More recent literature supports the view that inflation uncertainty can be approximated by the conditional variance of unanticipated shocks to inflation.[3]

In the literature most of previous studies analyzing the Friedman hypothesis have focused on the inflation and inflation uncertainty relationship (see for example Barnett *et al.*, 2020 and the references therein). However, despite the large number of empirical works there is still no consensus about either the direction or sign of this relationship. For instance, using a GARCH-type model Fountas and Karanasos (2007) find evidence of the positive effect of inflation on inflation uncertainty; see also Evans (1991), Holland (1995), Grier and Perry (1998), Fountas (2001), Apergis (2004), Kontonikas (2004), Daal *et al.* (2005). On the other side, the empirical investigation in Chang (2012) supports the view that inflation has a negative impact on inflation uncertainty during periods of high inflation volatility, and find no evidence of the impact of inflation on inflation uncertainty; see also Hwang (2001) and Wilson (2006).

Against this background, in this paper we try and reconcile two seemingly unrelated strands of literature by showing that the issue of inflation persistence and the inflation-nominal uncertainty link are closely related and should not be considered as separate matters. The paper builds on the work by Conrad and Karanasos (2015a) where it is shown that in time series models with in-mean and level effects there is a transmission of memory from the conditional variance to the conditional mean and vice versa. In particular, the authors consider the AR asymmetric power (AP) GARCH-in-mean-level (ML) model (the GARCH-M model was introduced by Engle et al.,1987; see also Conrad et al., 2010; Conrad and Karanasos, 2015b; and Karanasos and Zeng, 2013) and show that the model specification has an ARMA representation where the largest root of the AR part is closely linked not only to the inflation intrinsic persistence but to the persistence of the conditional variance of the process as well.

In the specific case of the inflation series, if the AR-APARCH-ML model can be used to capture the characteristic features of the process, then the inflation uncertainty parameter (i.e. the in-mean parameter in the conditional mean equation), induces a transmission of memory from the conditional variance to the conditional mean, that in turn affects the persistence properties of the level process. If this is the case, the result put and end to a long standing debate on the direction of causality of the inflation-nominal uncertainty relation showing that the two are closely related, therefore a shock to inflation uncertainty

[2] In the literature a positive relationship between inflation uncertainty and inflation has been supported by many studies, such as Apergis (2004), Wilson (2006) and Berument et al. (2009).

[3] See for example Grier *et al.* (2004), Fountas *et al.* (2006), Fountas and Karanasos, (2007), Chang *et al.* (2010), and Conrad and Karanasos (2015a,b)

affects the inflation inertia and vice versa. Some questions are however still open. For example, what are the implications for the predictability of inflation? In other words, how the forecasting properties of inflation would be affected by a shock to inflation uncertainty? Furthermore, inflation forecasts play an important role, since (i) policymakers react to forecasts due to inflation targeting adopted in most high income countries (see Clarida et al., 2000); and (ii) economic actors use inflation forecasts to decide upon future savings and expenditure levels. In this respect a possible shortcoming of the specification in Conrad and Karanasos (2015a) is that the model does not allow for structural breaks in inflation dynamics. Recent studies have shown that ignoring the presence of structural breaks can have important effects on the precision of inflation forecasting (see for example Evans and Wachtel, 1993; Berument et al. 2005, Caporale and Kontonikas, 2009; and Caporale et al. 2010). Accordingly, in this paper we build on Conrad and Karanasos (2015a) and consider a model where the conditional variance affects positively the conditional mean (in support of the Cukierman and Meltzer hypothesis) and the level has a positive impact on the conditional variance (in line with Friedman hypothesis), but the parameters of the model are allowed to change over time.

The contribution of this paper is twofold. First, we propose an AR-APGARCH-ML model with time varying coefficients that can be used to forecast inflation. Research over the past decade has documented considerable instability in inflation forecasting models.[4] However, most of this literature has focused on describing the evolution of macroeconomic dynamics. Empirical works that focus on the issue of modelling inflation rarely investigate the forecasting properties of the suggested model specifications. In general, studies that consider the forecasting ability of these models have been limited in both number and scope. In this respect, this work builds on the related literature by providing an extensive forecasting exercise comparing the suggested model to a number of specifications often used in empirical studies.

It will be clear from our analysis that predictability (forecasting) is closely interlinked with (first-order) persistence. Therefore we coin the term persistapredictability. In other words, higher persistence implies predictors with persistent structure. In particular, using the estimation results of the aforementioned model we compute various measures of first (and second)-order time varying persistence. Our work is close to Pivetta and Reis (2007) in spirit, however in this paper we depart from their study in an important way, that is we contribute to the measurement over time of inflation persistence by taking a different approach to the problem and estimating a model of inflation dynamics grounded in economic (rather than statistical) theory. In particular, we compute measures of persistence that not only distinguish between changes in the dynamics of inflation and its volatility (and their persistence), but also allow for feedback from volatility (inflation uncertainty) to the level of the process (inflation). Most importantly, there is a straightforward connection between persistence and predictability. One of the conditional measures of persistence is derived from the deterministic part of the optimal forecast, while another conditional criterion is retrieved from the forecast errors.

[4] See for example Stock and Watson (2007) or Stock and Watson (2009) for an excellent survey on the related literature.

In the related literature the issue of measuring persistence has been heavily investigated since the response to shocks of the economic system on inflation depends on the degree of persistence. Furthermore, the horizon of monetary policy actions should be targeted according to the persistence of inflation. However, the results of empirical works that document changes over time of inflation persistence in the United States are not conclusive. For example, Cogley and Sargent (2001, 2005) use a Bayesian state-space VAR model to model inflation dynamics and conclude that there was a change in the underlying characteristics of inflation reflecting a change in the structural characteristics of the economy and, possibly, a more active inflation targeting policy. In sharp contrast, Stock (2001) estimates the LAR using a rolling window estimation method and concludes that there is no indication of a marked decline in the persistence. A similar result is found in Pivetta and Reis (2007), where the LAR and the SAR are estimated using both Bayesian and rolling window estimation methods. The authors conclude that inflation persistence has been high in the United States and approximately unchanged over the entire post-war period. In this study, we argue that since time varying conditional volatility seems to be a characteristic feature of the inflation process in the United States (see for example Sensier and van Dijk, 2004) any suggested measure of inflation persistence should take into account the statistical properties of the conditional variance as well. In general, we believe that although the issue of inflation persistence has been heavily investigated over the last decade questions such as whether persistence has changed over time or remained constant, or whether it is structural or may vary according to specific monetary policy regimes are still open. The suggested measure of persistence might shed some light on these important issues.

The empirical results reveal several insights on the dynamics of inflation rate in the United States. First, we find evidence that the parameters in the models capturing inflation persistence change over time, in the conditional mean and/or the conditional variance. Therefore, not allowing for time varying coefficients in the estimation procedure would result in a less accurate modelling of the inflation process. Second, when it comes to forecasting, accounting for time varying parameters in the conditional mean equation improves the forecasting performance of the model. In general, comparing linear and nonlinear model specifications, we find that models that allows for time varying in-mean and level parameters have better forecasting performance than models that only allow for changes in the inflation autoregressive parameters. Third, using measures of persistence that only make use of autoregressive-type models to draw inference on the level persistence from the analysis of the estimated autoregressive coefficients may blur the picture of inflation dynamics since the implication of the suggested model is that higher level of uncertainty increases inflation persistence which in turn affects the forecasting properties of the process.

The outline of the paper is as follows. Section 2 introduces the time varying AR-APGARCH-ML model and presents its bivariate ARMA representation. In Section 3 we derive the first moment structure which is needed to obtain a new time varying measure of first-order persistence. Section 4 presents the estimation results along with the forecasting exercise for inflation series in the US. Section 5 compares the forecasting results of our model to other alternative linear and nonlinear model specifications. Section 6 presents the empirical results on the time varying persistence. Finally, Section 7 presents some concluding remarks.

2 The Model

In this section, we consider an AR(1)-APGARCH(1,1)-ML model, that is, a model in which there is bidirectional feedback between the conditional mean and variance, and two deterministic abrupt breaks (hereafter, DAB-AR(1;2)-ML model). In particular, we will examine the case of two breaks ($N = 2$) which occur at times $t-k_1$ and $t-k_2$ (with $k_2 > k_1$, $k_1 \in \mathbb{Z}_{>0}$ (the set of positive integers)); of course when $k_2 = k_1$ we have the case of one break), where the switch from one set of parameters to another is abrupt.

Let $\{y_t\}$, $t \in \mathbb{Z}$, be the inflation process, which follows a DAB-AR-M(1;2)-ML model:

$$y_t = \varphi(t) + \phi(t) y_{t-1} + \varsigma(t)\sigma_t^{\delta} + \varepsilon_t, \tag{1}$$

where $\varepsilon_t = e_t \sigma_t$, and the vector of the three deterministically varying coefficients, $\mathbf{m}(\tau)' = (\varphi(\tau), \phi(\tau), \varsigma(\tau))$ is given by

$$\mathbf{m}(\tau)' = \begin{cases} (\varphi_1, \phi_1, \varsigma_1) & \text{if} \quad \tau > t-k_1, \\ (\varphi_2, \phi_2, \varsigma_2) & \text{if} \quad t-k_2 < \tau \leq t-k_1, \\ (\varphi_3, \phi_3, \varsigma_3) & \text{if} \quad \tau \leq t-k_2, \end{cases}$$

with φ_n, ϕ_n, $\varsigma_n \in \mathbb{R}$ (the set of real numbers), $n = 1, 2, 3$, $\delta \in \mathbb{R}_{>0}$ (the set of positive real numbers), $\{e_t\}$ is a sequence of independent and identically distributed (*i.i.d*) random variables with zero mean and variance, $\mathbb{E}(e_t^2)$, and σ_t^2 is the conditional variance of y_t. [5] According to eq. (1) the breaks occur at times $t-k_1$ and $t-k_2$ and the switch from one set of parameters to another is abrupt. The time dependent autoregressive coefficient $\phi(t)$ naturally measures the intrinsic persistence in the level of y_t. By including σ_t^{δ} in the conditional mean we allow for feedback from the power transformed conditional variance of y_t to its level, captured by the deterministically varying in-mean coefficient $\varsigma(t)$. We denote the size of the breaks by $\Delta\phi_n = \phi_n - \phi_{n-1}$ and $\Delta\varsigma_n = \varsigma_n - \varsigma_{n-1}$, for $n = 2, 3$. For example, $\phi_2 = \phi_3 - \Delta\phi_3$ and $\phi_1 = \phi_3 - \Delta\phi_3 - \Delta\phi_2$.

The power transformed conditional variance, σ_t^{δ}, is positive with probability one and is a measurable function of $\mathcal{F}_{t-1}$, which in turn is the sigma-algebra generated by $\{y_{t-1}, y_{t-2}, \ldots\}$. We assume that σ_t^{δ} is specified as a time varying APGARCH-L(1,1) process:

$$\sigma_t^{\delta} = \omega(t) + \alpha(t) f(\varepsilon_{t-1}) + \beta(t)\sigma_{t-1}^{\delta} + d(t) y_{t-1}, \tag{2}$$

with

$$f(\varepsilon_{t-1}) = (|\varepsilon_{t-1}| - \gamma(t)\varepsilon_{t-1})^{\delta},$$

[5] Within the class of ARMA processes this specification is quite general and allows for intercept and slope shifts (see also Pesaran and Timmermann, 2005, Pesaran et al., 2006, and Koop and Potter, 2007).

where $|\gamma(t)| < 1$ for all t (for the APGARCH model with time invariant parameters see, for example, Ding *et al.*, 1993, and Karanasos and Kim, 2006). The following conditions are necessary and sufficient for $\sigma_t^\delta > 0$, for all t: $\omega(t) > 0$, $\alpha(t), \beta(t), d(t) \geq 0$, and $y_t \geq 0$, for all t. The vector of the five deterministically varying coefficients, $\mathbf{v}(\tau)' = (\omega(\tau), \alpha(\tau), \gamma(\tau), \beta(\tau), d(\tau))$ is given by

$$\mathbf{v}(\tau)' = \begin{cases} (\omega_1, \alpha_1, \gamma_1, \beta_1, d_1) & \text{if} \quad \tau > t - k_1, \\ (\omega_2, \alpha_2, \gamma_2, \beta_2, d_2) & \text{if} \quad t - k_2 < \tau \leq t - k_1, \\ (\omega_3, \alpha_3, \gamma_3, \beta_3, d_3) & \text{if} \quad \tau \leq t - k_2, \end{cases}$$

with $\omega_n, \alpha_n, \beta_n, d_n \in \mathbb{R}_{\geq 0}$ (the set of nonnegative real numbers), $|\gamma_n| < 1$, $n = 1, 2, 3$.

The model in eqs. (1) and (2) can be estimated by Quasi-Maximum Likelihood Estimation (QMLE) method. The asymptotic consistency of the QML estimator for the parametric GARCH-M model is established in Conrad and Mammen (2016).

Next we will introduce some important notation.

Notation 1 *i)We denote the time invariant r-th moment, $r \in \mathbb{Z}_{>0}$ (the set of positive integers) of the power transformed variance by $\mu_{rt} = \mathbb{E}(\sigma_t^{\delta r})$.*
ii) Similarly, $\kappa_r(t)$ denotes the r-th moment of $f(e_t)$: $\kappa_r(t) = \mathbb{E}[[f(e_t)]^r]$.

Clearly for $\delta \geq 1$, $\mu_{2/\delta,t} = \mathbb{E}(\sigma_t^2)$ is not a fractional moment only if δ is equal to 1 or 2. In all other cases $\mu_{2/\delta,t}$ has to be calculated numerically. However, if $\delta > 2$, the existence of the first moment, μ_{1t} guarantees that of $\mu_{2/\delta,t}$. Similarly, $\mu_{1+1/\delta,t} = \mathbb{E}(\sigma_t^{\delta+1})$ is not a fractional moment only if $\delta = 1/\lambda$ where $\lambda \in \mathbb{Z}_{>0}$. In all other cases $\mu_{1+1/\delta,t}$ has to be calculated numerically.

The APGARCH-L$(1,1)$ formulation in eq.(2) can readily be interpreted as having a weak time varying ARMA$(1,1)$ representation for the conditional variance:

$$\sigma_t^\delta = \omega(t) + c(t)\sigma_{t-1}^\delta + \alpha(t)v_{t-1} + d(t)y_{t-1}, \tag{3}$$

where

$$c(t) = \alpha(t)\kappa_1(t) + \beta(t), \quad \text{and} \quad v_t = f(\varepsilon_t) - \mathbb{E}[f(\varepsilon_t)\,|\mathcal{F}_{t-1}] = f(\varepsilon_t) - \kappa_1(t)\sigma_t^\delta,$$

and v_t is, by construction, an uncorrelated term with expected value 0. While the ε_t are the innovations to the level of y_t, the v_t can be considered the 'innovations' to the power transformed conditional variance of y_t.

By including the lagged y_t in the conditional variance equation (the so-called level effect) and σ_t^δ in the mean equation (the so-called in-mean effect), we allow for simultaneous feedback between the two variables.[6] Note that the parameter $c(t)$ measures the *intrinsic* memory or persistence in the conditional variance (see also Conrad and Karanasos, 2015a).

[6]Note that eq. (3) allows us to test for the Friedman hypothesis. Chang (2012), instead of a level effect considers an asymmetric GARCH specification for the conditional variance.

Next we will define the covariance matrix of the two 'shocks' ε_t and v_t, $\mathbf{\Sigma}_t$= $\mathbf{E}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t')$, where $\boldsymbol{\varepsilon}_t = (\varepsilon_t \; v_t)'$ and $\mathbf{E}(\cdot)$ denotes the elementwise expectation operator. First, we will denote the variances of the two 'shocks' and their covariance by

$$\sigma_{\varepsilon t} = \mathbb{E}(\varepsilon_t^2), \sigma_{vt} = \mathbb{E}(v_t^2), \sigma_{\varepsilon v,t} = \mathbb{E}(\varepsilon_t v_t).$$

The covariance matrix $\mathbf{\Sigma}_t$ is given by

$$\mathbf{\Sigma}_t = \begin{bmatrix} \sigma_{\varepsilon t} & \sigma_{\varepsilon v,t} \\ \sigma_{\varepsilon v,t} & \sigma_{vt} \end{bmatrix} = \begin{bmatrix} \mu_{2/\delta,t}\mathbb{E}(e_t^2) & \mu_{1+1/\delta,t}\widetilde{\kappa}(t) \\ \mu_{1+1/\delta,t}\widetilde{\kappa}(t) & \mu_{2t}\kappa(t) \end{bmatrix}, \tag{4}$$

where

$$\kappa(t) = \kappa_2(t) - \kappa_1^2(t), \; \widetilde{\kappa}(t) = \mathbb{E}[e_t f(e_t)].$$

In the following corollary we present expressions for $\kappa_r(t)$ and $\widetilde{\kappa}(t)$ under the assumption of Normality (see also Karanasos and Kim, 2006).

Corollary 1 *Consider the case where the term e_t is standard normal. Then $\mathbb{E}(e_t^2) = 1$, and $\kappa_r(t)$, $\widetilde{\kappa}(t)$ are given by*

$$\kappa_r(t) = \frac{1}{\sqrt{\pi}}\left[(1-\gamma(t))^{r\delta} + (1+\gamma(t))^{r\delta}\right] 2^{(\frac{r\delta}{2}-1)}\Gamma\left(\frac{r\delta+1}{2}\right),$$
$$\widetilde{\kappa}(t) = \frac{1}{\sqrt{2\pi}}\left[[1-\gamma(t)]^{\delta} - [1+\gamma(t)]^{\delta}\right]2^{(\delta/2)}\Gamma\left(\frac{\delta}{2}+1\right),$$

where $\Gamma(\cdot)$ is the Gamma function.
When $\delta = 1$ the above expressions reduce to $\widetilde{\kappa}(t) = -\gamma(t)$, $\kappa_1(t) = \sqrt{\frac{2}{\pi}}$, for all t, $\kappa_2(t) = 1 + \gamma^2(t)$ and therefore $\kappa(t) = \kappa_2(t) - \kappa_1^2(t) = 1 + \gamma^2(t) - \frac{2}{\pi}$, which implies that $\mathbf{\Sigma}_t$ becomes

$$\mathbf{\Sigma}_t = \mu_{2t}\begin{bmatrix} 1 & -\gamma(t) \\ -\gamma(t) & 1+\gamma^2(t)-\frac{2}{\pi} \end{bmatrix}. \tag{5}$$

Having defined the deterministically varying extension of the AR-APGARCH-ML model, next we will present its bivariate vector autoregressive moving average (BVARMA) formulation.

VARMA Formulation

To obtain the optimal predictors and the variance of y_t for the DAB-AR-ML model in eqs. (1) and (2) in the next proposition we will express eqs. (1) and (3) in a matrix form (the proof is straightforward).

Proposition 1 *eqs. (1) and (3) can be expressed in a matrix form as*

$$\mathbf{y}_\tau = \boldsymbol{\varphi}(\tau) + \mathbf{\Phi}(\tau)\mathbf{y}_{\tau-1} + \mathbf{J}\boldsymbol{\varepsilon}_\tau + \mathbf{Z}(\tau)\boldsymbol{\varepsilon}_{\tau-1}, \tag{6}$$

with $\tau \in \mathbb{Z}$, $\mathbf{y}_\tau = (y_\tau \; \sigma_\tau^\delta)'$, $\mathbf{J} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, where the three time varying coefficient matrices, $\boldsymbol{\varphi}(\tau)$, $\mathbf{\Phi}(\tau)$, and $\mathbf{Z}(\tau)$ are time invariant in each of the three segments:

$$\boldsymbol{\varphi}_n = \begin{bmatrix} \varphi_n + \varsigma_n\omega_n \\ \omega_n \end{bmatrix}, \mathbf{\Phi}_n = \begin{bmatrix} \phi_n + \varsigma_n d_n & \varsigma_n c_n \\ d_n & c_n \end{bmatrix}, \mathbf{Z}_n = \begin{bmatrix} 0 & \varsigma_n\alpha_n \\ 0 & \alpha_n \end{bmatrix}, \begin{cases} n=1 & \text{if} \quad \tau > t-k_1, \\ n=2 & \text{if} \quad t-k_2 < \tau \le t-k_1, \\ n=3 & \text{if} \quad \tau \le t-k_2. \end{cases}$$

For notational convenience we will interchangeably use $\mathbf{\Phi}_3$ or $\mathbf{\Phi}$ and $\mathbf{Z}_3$ or $\mathbf{Z}$. We will term the deterministically varying bivariate expression in eq. (6) the DAB-BVARMA$(1,1;2)$ representation.[7]

In what follows we will employ the above representation to derive explicit formulas for the optimal predictors (and the variances) of y_t and σ_t^δ in eqs. (1) and (2), respectively.[8] These are needed in order to obtain time varying first and second-order measures of persistence.

3 Persistapredictability

3.1 Optimal Forecasts

In this section we provide an equivalent explicit solution representation of the DAB-BVARMA$(1,1;2)$ process in eq. (6), which generates explicit formulas for the optimal predictors and the bidimensional time varying covariance matrix of $\{\mathbf{y}_\tau\}$, $\tau = t + r$, $r \in \mathbb{Z}_{\geq 0}$ (the set of nonnegative integers).

First, let $\lambda_{\max}(\mathbf{X})$ denote the modulus of the largest eigenvalue of $\mathbf{X}$. The following theorem holds (the proof is presented in the Appendix 1).

Theorem 1 *An equivalent explicit solution representation of the bivariate system in eq. (6), subject to the initial condition $\mathbf{y}_{\tau-k}$, for $k \geq k_2 + r$, is given by*

$$\mathbf{y}_{\tau,k} = \mathbf{E}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}) + \mathbf{FE}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}), \tag{7}$$

where

$$\mathbf{E}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}) = \boldsymbol{\varphi}_k(\tau) + \mathbf{\Phi}_1^{k_1+r}\mathbf{\Phi}_2^{k_2-k_1}\mathbf{\Phi}^{k-k_2-1}(\mathbf{\Phi}\mathbf{y}_{\tau-k} + \mathbf{Z}\boldsymbol{\varepsilon}_{\tau-k}), \tag{8}$$

$$\begin{aligned}\mathbf{FE}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}) = {} & \mathbf{J}\boldsymbol{\varepsilon}_\tau + \sum_{\ell=1}^{k_1+r}\mathbf{\Phi}_1^{\ell-1}(\mathbf{\Phi}_1\mathbf{J} + \mathbf{Z}_1)\boldsymbol{\varepsilon}_{\tau-\ell} + \mathbf{\Phi}_1^{k_1+r}\{\sum_{\ell=1}^{k_2-k_1}\mathbf{\Phi}_2^{\ell-1}(\mathbf{\Phi}_2\mathbf{J} + \mathbf{Z}_2)\boldsymbol{\varepsilon}_{t-k_1-\ell} \\ & + \mathbf{\Phi}_2^{k_2-k_1}[\sum_{\ell=1}^{k-k_2-1}\mathbf{\Phi}^{\ell-1}(\mathbf{\Phi}\mathbf{J} + \mathbf{Z})\boldsymbol{\varepsilon}_{t-k_2-\ell}]\},\end{aligned} \tag{9}$$

and if $\lambda_{\max}(\mathbf{\Phi}_n) \neq 1$, $n = 1, 2, 3$, then

$$\boldsymbol{\varphi}_k(\tau) = (\mathbf{I} - \mathbf{\Phi}_1^{k_1+r})(\mathbf{I} - \mathbf{\Phi}_1)^{-1}\boldsymbol{\varphi}_1 + \mathbf{\Phi}_1^{k_1+r}[(\mathbf{I} - \mathbf{\Phi}_2^{k_2-k_1})(\mathbf{I} - \mathbf{\Phi}_2)^{-1}\boldsymbol{\varphi}_2 + \mathbf{\Phi}_2^{k_2-k_1}(\mathbf{I} - \mathbf{\Phi}^{(k-k_2)})(\mathbf{I} - \mathbf{\Phi})^{-1}\boldsymbol{\varphi}]. \tag{10}$$

In the above expression if $\lambda_{\max}(\mathbf{\Phi}_n) = 1$, then $(\mathbf{I} - \mathbf{\Phi}_n^{k_n-k_{n-1}})(\mathbf{I} - \mathbf{\Phi}_n)^{-1}$, with $k_0 = -r$ and $k_3 = k$, should be replaced by $\sum_{\ell=0}^{k_n-k_{n-1}-1}\mathbf{\Phi}_n^\ell$ (a similar argument holds for any of the analogous cases that follow).

[7] As pointed out by Conrad and Karanasos (2015a) the AR(1)-APGARCH(1,1)-M model with constant coefficients is observationally equivalent to an ARMA$(2,1)$ process with the largest autoregressive root close to one if ϕ is close to one or c is close to one (or both). Clearly, if $\phi = 0$, $c = 1$ and there are no breaks the reduced form representation of the AR-M specification coincides with the IMA$(1,1)$ model proposed by Stock and Watson (2007).

[8] Notice that, as pointed ouy by Pivetta and Reis (2007), including other variables would lead to an assessment of predictability. Since here we focus on persistence and predictability (the two are interlinked), we work with a univariate GARCH-ML model.

The above theorem expresses the explicit solution representation, $\mathbf{y}_{\tau,k}$, in terms of the $(k+r)$-step ahead optimal in (L_2 sense) linear predictor, $\mathbf{E}(\mathbf{y}_\tau | \mathcal{F}_{\tau-k})$, and the associated forecast error, $\mathbf{FE}(\mathbf{y}_\tau | \mathcal{F}_{\tau-k})$. Clearly, if $k_2 = k_1$ eq. (7) gives the solution in the case of one break, whereas if $k_2 = k_1 = k$, it gives the general solution when there is no time variation. For example, for the time invariant case, since $\mathbf{\Phi}_1 = \mathbf{\Phi}_2 = \mathbf{\Phi}$ and $\mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{Z}$, the forecast error in eq. (7) reduces to

$$\mathbf{FE}(\mathbf{y}_\tau | \mathcal{F}_{\tau-k}) = \mathbf{J}\boldsymbol{\varepsilon}_\tau + \sum_{\ell=1}^{k+r-1} \mathbf{\Phi}^{\ell-1}(\mathbf{\Phi}\mathbf{J} + \mathbf{Z})\boldsymbol{\varepsilon}_{\tau-\ell}. \tag{11}$$

Accordingly the general solutions when $k \leq k_1 + r$ and $k_1 + r < k < k_2 + r$ can be obtained along the lines of Theorem 1 and are equivalent to the time invariant case and the case when there is one break, respectively.

In this section, in the context of the AR-APGARCH-ML model, we show the importance of taking into account abrupt breaks for the in-sample forecasting. Having found an explicit formula for the general solution of the DAB-BVARMA$(1,1;2)$ representation, in the next section we will show how these results can be used to derive the first moment structure.

3.2 First Moment Structure

In the sequel we will use the notation

$$\begin{aligned}
\mathbf{E}(\mathbf{y}_\tau) &= \lim_{k\to\infty} \mathbf{E}(\mathbf{y}_\tau | \mathcal{F}_{\tau-k}), \\
\mathbf{FE}(\mathbf{y}_\tau) &= \lim_{k\to\infty} \mathbf{FE}(\mathbf{y}_\tau | \mathcal{F}_{\tau-k}), \\
\boldsymbol{\varphi}(\tau) &= \lim_{k\to\infty} \boldsymbol{\varphi}_k(\tau).
\end{aligned}$$

Assumption 1 (First-Order). We assume that $\lambda_{\max}(\mathbf{\Phi}_n) < 1$, $n = 1, 3$.

Proposition 2 *Let Assumption 1 hold. The expected value of the bivariate system in eq. (6),* $\mathbf{E}(\mathbf{y}_\tau)$*, is equal to* $\boldsymbol{\varphi}(\tau)$*, which is given by*

$$\boldsymbol{\varphi}(\tau) = (\mathbf{I} - \mathbf{\Phi}_1^{k_1+r})(\mathbf{I} - \mathbf{\Phi}_1)^{-1}\boldsymbol{\varphi}_1 + \mathbf{\Phi}_1^{k_1+r}[(\mathbf{I} - \mathbf{\Phi}_2^{k_2-k_1})(\mathbf{I} - \mathbf{\Phi}_2)^{-1}\boldsymbol{\varphi}_2 + \mathbf{\Phi}_2^{k_2-k_1}(\mathbf{I} - \mathbf{\Phi})^{-1}\boldsymbol{\varphi}]. \tag{12}$$

Notice that when $r \to \infty$*, the expected value (under Assumption 1) is equivalent to the one for the invariant case:* $\boldsymbol{\varphi}(\tau) = (\mathbf{I} - \mathbf{\Phi}_1)^{-1}\boldsymbol{\varphi}_1$*. Similarly,* $\mathbf{FE}(\mathbf{y}_\tau)$ *is given by*

$$\begin{aligned}
\mathbf{FE}(\mathbf{y}_\tau) = \mathbf{J}\boldsymbol{\varepsilon}_\tau + \sum_{\ell=1}^{k_1+r} \mathbf{\Phi}_1^{\ell-1}(\mathbf{\Phi}_1\mathbf{J} + \mathbf{Z}_1)\boldsymbol{\varepsilon}_{\tau-\ell} + \mathbf{\Phi}_1^{k_1+r}\{ \sum_{\ell=1}^{k_2-k_1} \mathbf{\Phi}_2^{\ell-1}(\mathbf{\Phi}_2\mathbf{J} + \mathbf{Z}_2)\boldsymbol{\varepsilon}_{t-k_1-\ell} \\
+ \mathbf{\Phi}_2^{k_2-k_1}[\sum_{\ell=1}^{\infty} \mathbf{\Phi}^{\ell-1}(\mathbf{\Phi}\mathbf{J} + \mathbf{Z})\boldsymbol{\varepsilon}_{t-k_2-\ell}]\}.
\end{aligned}$$

Proposition 3 *Let Assumption 1 hold. The Wold-Cramér representation of the bivariate system in eq. (6),* $\mathbf{y}_\tau = \lim_{k\to\infty} \mathbf{y}_{\tau,k}$*, is given by*

$$\mathbf{y}_\tau = \boldsymbol{\varphi}(\tau) + \mathbf{FE}(\mathbf{y}_\tau). \tag{13}$$

The solution of eq. (6) in eq. (13) is decomposed in two orthogonal parts, a deterministic part which is the unconditional mean, and a zero random part, that is the limit of the forecast errors.

In the next section we will show how the results on predictability and the first moment structure can be used to derive a time varying first-order measure of persistence. In the Appendix 1 we will derive an explicit formula for the bidimensional time varying covariance matrix of $\{y_t\}$, which, as noted above, is needed in order to obtain a time varying measure of second-order persistence (see the Appendix 3 and 4).

3.3 Time Varying Persistence

The most often applied time invariant measures of first-order (or mean) persistence are the LAR (largest AR root), and the SAR (sum of the AR coefficients). As pointed out by Pivetta and Reis (2007) in relation to the issue of recidivism by monetary policy its occurrence depends very much on the model used to test the natural rate hypothesis, i.e., the hypothesis that the SAR or the LAR for inflation data is equal to one. Obviously, both measures would ignore the presence of breaks, in-mean and possible level effects and, hence, potentially under or over estimate the persistence in the levels, which is partly induced by its biderictenional feedback with the persistence in the conditional variance.

The LAR has been used to measure persistence in the context of testing for the presence of unit roots (see, for details, Pivetta and Reis, 2007). The authors find no evidence pointing to a rejection of a unit root in inflation. However, as we show in Canepa *et al.* (2019) if the in-mean mechanism together with the possible presence of breaks in the in-mean parameter are ignored, then conventional procedures (such as unit root tests) for estimating the persistence in the mean may lead to biased estimates. In particular, they might falsely indicate a unit root, and, hence, suggest the modeling of the differenced series rather than their levels. Fiorentini and Sentana (1998) argue that any reasonable measure of shock persistence should be based on the IRFs (impulse response functions). Therefore we will consider two more alternative measures of first-order persistence (four in total):

- 1st. LAR
- 2nd. 1/(1-SAR)
- 3rd. The unconditional mean
- 4th. The sum of the MA coefficients in the Wold representation

It will be clear from the following analysis that predictability (forecasting) is closely interlinked with (first-order) persistence. Therefore as mentioned above we coin the term *persistapredictability.* In other words, higher first-order persistence implies predictors with persistent structure.

Below we will show how the four aforementioned measures of first-order inflation persistence for the time invariant univariate case (the model without in-mean and level effects and breaks) should be modified in order to become applicable to the general time varying bivariate case.

Next we will use the notation $\mathbf{j} = (1\ 1)'$. Setting $\boldsymbol{\varphi}_1 = \boldsymbol{\varphi}_2 = \boldsymbol{\varphi} = \mathbf{j}$ in eq. (10), that is assuming unit mean and volatility (BVARMA) drifts, we obtain the vector $\mathbf{P}^{(AR)}_{\tau,k}\mathbf{j}$, where $\mathbf{P}^{(AR)}_{\tau,k}$ is given by

$$\mathbf{P}^{(AR)}_{\tau,k} = (\mathbf{I}-\mathbf{\Phi}_1^{k_1+r})(\mathbf{I}-\mathbf{\Phi}_1)^{-1} + \mathbf{\Phi}_1^{k_1+r}[(\mathbf{I}-\mathbf{\Phi}_2^{k_2-k_1})(\mathbf{I}-\mathbf{\Phi}_2)^{-1} + \mathbf{\Phi}_2^{k_2-k_1}(\mathbf{I}-\mathbf{\Phi}^{(k-k_2)})(\mathbf{I}-\mathbf{\Phi})^{-1}]. \quad (14)$$

The limit of $\mathbf{P}^{(AR)}_{\tau,k}$ as $k \to \infty$, denoted hereafter by $\mathbf{P}^{(AR)}_{\tau}$, under Assumption 1, is given by

$$\mathbf{P}^{(AR)}_{\tau} = \mathbf{P}^{(AR)}_{\tau,k} + \mathbf{\Phi}_1^{k_1+r}\mathbf{\Phi}_2^{k_2-k_1}\mathbf{\Phi}^{(k-k_2)}(\mathbf{I}-\mathbf{\Phi})^{-1}. \quad (15)$$

Notation 2 *i) We will denote the two elements in the first row of* $\mathbf{P}^{(AR)}_{\tau,k}$ *in eq. (14) by* $P^{(AR)}_k(y_\tau\,|\varepsilon)$ *and* $P^{(AR)}_k(y_\tau\,|v)$*, and their sum by* $P^{(AR)}_k(y_\tau)$*.*
ii) The limits of $P^{(AR)}_k(y_\tau\,|\varepsilon)$ *and* $P^{(AR)}_k(y_\tau\,|v)$ *as* $k \to \infty$*, that is the two elements in the first row of* $\mathbf{P}^{(AR)}_{\tau}$ *in eq. (15), will be denoted by* $P^{(AR)}(y_\tau\,|\varepsilon)$ *and* $P^{(AR)}(y_\tau\,|v)$*, respectively, and their sum by* $P^{(AR)}(y_\tau)$*.*

Setting $\boldsymbol{\varepsilon}_\tau = \mathbf{j}$ for all τ in eq. (9), that is assuming unit mean and volatility shocks, we obtain the vector $\mathbf{P}^{(MA)}_{\tau,k}\mathbf{j}$, where $\mathbf{P}^{(MA)}_{\tau,k}$ is given by

$$\begin{aligned}\mathbf{P}^{(MA)}_{\tau,k} \;=\;& \mathbf{J}+(\mathbf{I}-\mathbf{\Phi}_1^{k_1+r})(\mathbf{I}-\mathbf{\Phi}_1)^{-1}(\mathbf{\Phi}_1\mathbf{J}+\mathbf{Z}_1) + \mathbf{\Phi}_1^{k_1+r}[(\mathbf{I}-\mathbf{\Phi}_2^{k_2-k_1})(\mathbf{I}-\mathbf{\Phi}_2)^{-1}(\mathbf{\Phi}_2\mathbf{J}+\mathbf{Z}_2)\\ &+\mathbf{\Phi}_2^{k_2-k_1}(\mathbf{I}-\mathbf{\Phi}^{(k-k_2-1)})(\mathbf{I}-\mathbf{\Phi})^{-1}(\mathbf{\Phi}\mathbf{J}+\mathbf{Z})]. \quad (16)\end{aligned}$$

Notice that if in the above equation we set $\mathbf{J} = \mathbf{I}$ and $\mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{Z} = \mathbf{0}$, that is removing the appearance in the mean of the moving average term (present due to the in-mean effects), then it can be shown, using straightforward matrix algebra, that $\mathbf{P}^{(MA)}_{\tau,k}$ reduces to $\mathbf{P}^{(AR)}_{\tau,k}$. The limit of $\mathbf{P}^{(MA)}_{\tau,k}$ as $k \to \infty$, denoted hereafter by $\mathbf{P}^{(MA)}_{\tau}$, under Assumption 1, is given by

$$\mathbf{P}^{(MA)}_{\tau} = \mathbf{P}^{(MA)}_{\tau,k} + \mathbf{\Phi}_1^{k_1+r}\mathbf{\Phi}_2^{k_2-k_1}\mathbf{\Phi}^{(k-k_2-1)}(\mathbf{I}-\mathbf{\Phi})^{-1}(\mathbf{\Phi}\mathbf{J}+\mathbf{Z}). \quad (17)$$

Notation 3 *i) We will denote the two elements in the first row of* $\mathbf{P}^{(MA)}_{\tau,k}$ *in eq. (16) by* $P^{(MA)}_k(y_\tau\,|\varepsilon)$ *and* $P^{(MA)}_k(y_\tau\,|v)$*, and their sum by* $P^{(AR)}_k(y_\tau)$
ii) The limits of $P^{(MA)}_k(y_\tau\,|\varepsilon)$ *and* $P^{(MA)}_k(y_\tau\,|v)$ *as* $k \to \infty$*, that is the two elements in the first row of* $\mathbf{P}^{(MA)}_{\tau}$ *in eq. (17), will be denoted by* $P^{(MA)}(y_\tau\,|\varepsilon)$ *and* $P^{(MA)}(y_\tau\,|v)$*, respectively, and their sum by* $P^{(MA)}(y_\tau)$*.*

The time varying versions of the four first-order (or mean) persistence measures that are able to take into account the presence of breaks and to distinguish between the effects of a *mean shock* and a *volatility shock* on the level and conditional variance, respectively, are as follows:

- 1st. $\max\left(\lambda_{\max}(\mathbf{\Phi}), \lambda_{\max}(\mathbf{\Phi}_1)\right)$

(the equivalent of the LAR where $\lambda_{\max}(\cdot)$ has been defined in Section 3.1)

- 2nd. $P^{(AR)}(y_\tau)$

(the corresponding of the 1/1-SAR, see eq. (15) and Notation 2(ii))

- 3rd. $\mathbb{E}(y_t)$

(the unconditional mean, that is the first element of $\boldsymbol{\varphi}(\tau)$ in eq. (12))

- 4th. $P^{(MA)}(y_\tau)$

(the analogous to the sum of the Green functions[9], which is retrieved from the sum of the Green Matrices (SGM) in eq. (17), see also Notation 3(ii))

The first measure is identical for both the inflation and its conditional variance and it prohibits time variation. The second one excludes the participation of the drifts and the presence (due to the in-mean effect) of the MA terms. The third metric ignores the presence of the MA structure. The fourth criterion incorporates the involvement of the MA terms but it does not take into account the drifts.

Notice that the last three measures are unconditional ones. [10] The corresponding conditional 2nd and 4th measures are given by $P_k^{(AR)}(y_\tau)$ and $P_k^{(MA)}(y_\tau)$, see Notations 2(i) and 3(i), respectively. The conditional analogous of the mean is the first element of $\boldsymbol{\varphi}_k(\tau)$ in eq. (10).

There is a straightforward connection between persistence and predictability. The second conditional measure of persistence, $P_k^{(AR)}(y_\tau)$, is derived from the deterministic part of the optimal forecast, that is $\boldsymbol{\varphi}_k(\tau)$ in eq. (10). The fourth conditional criterion, $P_k^{(MA)}(y_\tau)$, is retrieved from the forecast errors in eq. (9).

Time Invariant Case

Next we will see how the forenamed measures of persistence simplify when the coefficients are constant. The limits of $\boldsymbol{\varphi}(\tau)$, $\mathbf{P}_\tau^{(AR)}$, $\mathbf{P}_\tau^{(MA)}$ (see eqs. (12), (15) and (17), respectively) as $r \to \infty$, denoted hereafter by $\boldsymbol{\varphi}$, $\mathbf{P}^{(AR)}$, $\mathbf{P}^{(MA)}$ respectively, are given by (setting $\boldsymbol{\varphi} = \boldsymbol{\varphi}_1$, $\mathbf{\Phi} = \mathbf{\Phi}_1$ and $\mathbf{Z} = \mathbf{Z}_1$):

$$\begin{aligned} \mathbf{P}^{(AR)} &= (\mathbf{I} - \mathbf{\Phi})^{-1}, \\ \boldsymbol{\varphi}(\tau) &= (\mathbf{I} - \mathbf{\Phi})^{-1}\boldsymbol{\varphi}, \\ \mathbf{P}^{(MA)} &= \mathbf{J} + (\mathbf{I} - \mathbf{\Phi})^{-1}(\mathbf{\Phi}\mathbf{J} + \mathbf{Z}). \end{aligned} \tag{18}$$

[9] The time varying MA coefficients in the Wold-Cramér representation are called Green functions (see, for example, Karanasos *et al.*, 2022).

[10] Moreover, to save space the equivalent persistence measures for the power transformed conditional variance are not reported but are available upon request.

Next we will use the notation: $D = (1-\phi)(1-c) - d\varsigma$ and $\max(a \pm b) = \max(a+b, a-b)$. The four (unconditional) measures of persistence for the time invariant case are obtainable from the eqs. (18) by straightforward algebra:

- 1st. $\max(\frac{1}{2}[(\phi + c + \varsigma d) \pm \sqrt{(\phi - c + \varsigma d)^2 + 4d\zeta c})$
- 2nd. $P^{(AR)}(y) = \frac{1-c(1-\varsigma)}{D}$
- 3rd. $\mathbb{E}(y_t) = \frac{1}{D}[(1-c)\varphi + \varsigma\omega]$
- 4th. $P^{(MA)}(y_\tau) = 1 + \frac{(1-c)\phi + (d+\alpha)\varsigma}{D}$

It is apparent that if $\varsigma, d > 0$ (positive in-mean and level effects) then higher either in-mean or level effects increase the persistence. For example, regarding the LAR we have:

$$\underbrace{\max(\frac{1}{2}[(\phi + c + \varsigma d) \pm \sqrt{(\phi - c + \varsigma d)^2 + 4d\zeta c})}_{\text{positive in-mean and level effects}} > \underbrace{\max(\phi, c)}_{\text{in-mean effects}} \geq \underbrace{\phi}_{\text{AR model}}.$$

When there are no level effects the measures simplify to:

- 1st. $\max(\phi, c)$
- 2nd. $P^{(AR)}(y_\tau) = \frac{1}{1-\phi}(1 + \frac{1+c\varsigma}{1-c})$
- 3rd. $\mathbb{E}(y_t) = \frac{1}{1-\phi}(\varphi + \frac{\varsigma\omega}{1-c})$
- 4th. $P^{(MA)}(y_\tau) = 1 + \frac{1}{1-\phi}(\phi + \frac{\varsigma\alpha}{1-c})$

It is evident that $\varsigma > 0$ ($\varsigma < 0$) increase (decrease) the persistence. Setting $\varsigma = 0$, we obtain the well known measures of persistence for the AR(1) model:

$$\text{1st. } LAR = \phi, \text{ 2nd and 4th. } 1/(1-SAR) = \frac{1}{1-\phi}, \text{ 3rd. } \mathbb{E}(y_t) = \frac{\varphi}{1-\phi}.$$

4 Data and Empirical Results

In our empirical application we consider log-differences of quarterly data of Personal Consumption Expenditure (CPE) in the United States from 1947Q1 to 2021Q3. The data were collected from the Federal Reserve of St. Louis data bank.

4.1 Testing for Breaks

The first step in the estimation procedure is to identify possible points of parameter changes. Failure to identify in-sample breaks that change the data generating process of the inflation series produces biased parameter estimates and affects the model's out-of-sample forecasting performance. Ideally, if information on breaks, such as breakpoints and break sizes, is known, we can decide the estimation window size according to the trade off between the bias and forecast error variance to improve the out-of-sample forecasting errors (see Pesaran and Timmermann, 2005).

With this target in mind the Perron *et al.* (2020) sequential test on inflation series is used to identify possible breaks during the sample period. The authors provide a comprehensive treatment for testing jointly for structural changes not only in the conditional mean but also in the conditional variance. The inference procedure is more general than the widely used Bai and Perron (2003) as it allows for the break dates in the conditional mean and the conditional variance to be different or overlap.[11] This property of the test constitutes an important step forward since a growing number of empirical works have found that regime changes in the conditional variance are a characteristic feature of many macroeconomic time series and that regime shifts often do not occur at the same time as regime shifts in the conditional mean (see Sensier and Dijk, 2004 and the references therein). In addition the test is particularly suitable for the model in Eq. (1) since the procedure only requires mild assumptions on the innovations and conditional heteroskedasticity is permitted (see Perron *et al.*, 2020 for more details).

To locate the breaks we estimate a model for the inflation series with only a constant. The resulting specification is a pure structural change model that allows for k_m breaks in the conditional mean parameters and k_v breaks (or $k_v + 1$ regimes) in the variance of the innovations occurring at unknown dates.[12] The breaks in the variance and in the conditional mean coefficients can happen at different times, hence the k_m-vector breaks in the conditional mean and k_v-vector breaks in variance can have all distinct elements or they can overlap partly or completely.

Following Perron *et al.* (2020) we adopt a specific to general procedure that uses the sequential test to determine the number of breaks in the conditional mean and variance allowing for a given number of breaks in the other component. The test evaluates whether including an additional break is warranted.

Let $(T_1^m, ..., T_{k_m}^m; T_1^v, ..., T_{k_v}^v)$ be the estimates of the break dates in the conditional mean parameter and the error variance obtained jointly by maximizing the quasi-likelihood function ($\log L_T$) assuming k_m breaks in conditional mean and k_v breaks in variance. To test whether an additional break in conditional mean constant is present we use

$$\begin{aligned} \sup Seq_T\left(k_m+1, k_v | k_m, k_v\right) \quad = \quad & \max_{1 \le j \le k_m+1} \sup_{\tau \in \Lambda_{j,l}^m} \log L_T\left(T_1^m, ..., T_{k_m}^m, \tau, T_1^v, ..., T_{k_v}^v\right) \\ & - \log L_T\left(T_1^m, ..., T_{k_v}^m, \tau, T_1^v, ..., T_{k_v}^v\right) \end{aligned} \tag{19}$$

[11] Note the issue of testing for structural breaks in the context of forecasting has been heavily investigated in the literature (see for example Peseran and Timmerman, 2005). For example, Maheu and Gordon (2008) suggest a Bayesian method for forecasting under structural breaks and assume that the post-break distribution is given by a subjective prior. Similarly, Peseran *et al.* (2006) propose a Bayesian estimation and prediction procedure that allows for the possibility of new breaks over the forecast horizon, taking account of the size and duration of past breaks (if any) by means of a hierarchical hidden Markov chain model. Predictions are formed by integrating over the hyper parameters from the meta distributions that characterize the stochastic break point process.

[12] Note that Perron *et al.* (2020) consider a model specification with martingale difference errors and not the GARCH-in mean type of model considered in this paper. However, given that the GARCH-in mean only involves including the estimated standard deviation in the conditional mean equation the limit theory provide in Perron *et al.* (2020) might also be valid for the type of model considered in this paper. Providing the limit distribution of the Perron *et al.* (2020) test for the GARCH-in mean case is outside the scope of the paper.

where $\Lambda_{j,u}^{m}=\left\{\tau,T_{j-1}^{m}+\left(T_{j}^{m}-T_{j-1}^{m}\right)u\leq\tau\leq T_{j-1}^{m}-\left(T_{j}^{m}-T_{j-1}^{m}\right)u\right\}$ and u is a truncation imposing a minimal length for each segment. This amounts to performing k_m+1 tests for a single break in the conditional mean for each of the k_m+1 regimes defined by the partition $\left\{T_1^m,...,T_{k_m}^m\right\}$.

Note that there are different scenarios when allowing breaks in the conditional mean and in the variance to happen at different dates, since $\left(T_1^m,...,T_{k_m}^m\right)$ and $\left(T_1^v,...,T_{k_v}^v\right)$ can partly or completely overlap or be altogether different.

The computation of the estimates of the breaks using a sequential procedure is obtained by QML and the limit distribution of the test for martingale difference errors is provided by Perron *et al.* (2020).

Similarly, to test whether an additional break in the variance is present the test statistic is given by

$$\begin{aligned}\sup Seq_T\left(k_m,k_v+1|k_v,k_v\right) &= (2/\varrho)\max_{1\leq j\leq k_v+1}\sup_{\tau\in\Lambda_{j,l}^{m}}\log L_T\left(T_1^m,...,T_{k_m}^m,\tau,T_1^v,...,T_{k_v}^v\right) \\ &\quad -\log L_T\left(T_1^m,...,T_{k_m}^m,\tau,T_1^v,...,T_{k_v}^v\right)\end{aligned} \tag{20}$$

where $\Lambda_{i,u}^{m}=\left\{\tau,T_{i-1}^{v}+\left(T_{i}^{v}-T_{i-1}^{v}\right)u\leq\tau\leq T_{j-1}^{v}-\left(T_{i}^{v}-T_{i-1}^{v}\right)u\right\}$ and $(2/\varrho)$ is a factor that is needed to ensure that the limit distribution of the test is free of nuisance parameters.

As in the previous case the test is applied to each segment containing the T_{k_v-1} to T_{k_v} $(i=1,...,k_v+1)$. In particular, the procedure involves using a sequence of tests, where the conclusion of a rejection in favour of a model with (k_v+1) breaks if the overall minimal value of the sum of squared residuals is sufficiently smaller than the sum of the squared residuals from the k_v break model. Also in this case, the computation of the estimates of the breaks occurring at unknown date is obtained by QMLE and the limit distribution of the test is given in Perron *et al.* (2020).

The results of the structural break test are reported in Table 1. In the top panel of Table 1 the second, third and fourth columns report the calculated value of the statistic for the hypotheses of $k_m=0,...,3$ breaks in the conditional mean given $k_v=0,...,3$ breaks in the variance. Similarly, in the bottom part of Table 1, the second, third and fourth columns report the calculated value of statistic for the hypothesis of $k_v=0,...,3$ breaks in the variance given $k_m=0,...,3$ breaks in the conditional mean. We assume that there are at most three breaks in the conditional mean and/or in the variance. Finally, in the last column of Table 2 the estimated break dates are reported.

Table 1. Perron *et al.* (2020) sequential procedure to determine breaks in the inflation data.

		sup $Seq_{m,T}$		
	$k_m = 1$	$k_m = 2$	$k_m = 3$	Break Dates
$k_v = 0$	7.04	14.53**	13.99**	1990Q4
$k_v = 1$	7.04	14.50**	14.53**	1990Q4
$k_v = 2$	4.53	14.49	14.73	
$k_v = 3$	5.56	13.58**	13.31**	1990Q4
		sup $Seq_{v,T}$		
	$k_v = 1$	$k_v = 2$	$k_v = 3$	Break Dates
$k_m = 0$	10.62**	11.25*	11.38*	1990Q4
$k_m = 1$	8.42	11.23*	15.91***	
$k_m = 2$	11.25**	3.15	11.71*	
$k_m = 3$	5.83	3.10	11.70*	1954Q4

Note: Table 1 reports the calculated Perron et al. (2020) test for structural breaks. Note ***), **), *) indicates rejection of the null hypothesis at $1\%, 5\%$ and 10%, respectively.

Looking at the calculated values of the test, the sequential procedure Sup $\text{Seq}_{m,T}$ detects three breaks in the mean in 1990Q4, irrespective of how many variance breaks are accounted for. On the other side, the sequential procedure using the Sup $\text{Seq}_{v,T}$ test detects one break in the error variance that also occurred in 1990Q4 and a break in 1954Q4 in the conditional variance only.

The first break took place in the 1950s when the economy experienced a surge in inflation in the run-up to the Korean war. In particular, the period of 1950-1953 coincides with mobilization for the Korean war, and the recession the followed after the end of the war. The second break occurred at the beginning of the 1990s. This finding is interesting since it supports the empirical results in Del Negro *et al.* (2020) where it is found that the Phillips curve flattened around the 1990s. The authors argue that around that period the path of inflation and unemployment changed mainly due to three reasons. First, the labour markets undertook major changes making unemployment a poorer indicator of both the degree of resource spare capacity in the economy and of the cost pressures faced by firms. Second, firms' pricing decisions became less sensitive to these cost pressures. Third, monetary policy became more successful in stabilizing inflation using for example inflation targeting policies (see also McLeay and Tenreyro, 2019). The view that the stability of inflation over the past 30 years was not due to cyclical behavior but a major structural change is also corroborated in Ball and Mazumder (2011), Bobeica and Jarocinski (2017), Eser *et al.* (2020).

Accordingly, below we estimate the DAB-AR(1; 2)-ML model in eq. (1), allowing for persistence of inflation (as captured by the autoregressive coefficient, $\phi(t)$) and the in-mean coefficient, $\varsigma(t)$, to switch across breakpoints. Similarly, we allow for the coefficients of eq. (2) to switch across the break points.

4.2 Estimated Models

Table 2 reports the estimated parameters of the model in eqs.(1)-(2) for the inflation series and the relative misspecification tests. In particular, the top part of Table 2 reports the estimated parameters for the conditional mean, whereas the coefficients for the conditional variance are given in Panel B. Note that in the preliminary model selection procedure a number of specifications were estimated, however, in Table 2 we only report the subset of best model specifications. All the models were estimated with $\delta = 1$ in eqs. (1)-(2).[13]

In Table 2 four model specifications are reported, we label these models M1, M2, M3, M4, respectively. Model M1 is a simple AR(1)-(P)GARCH$(1,1)$-M specification with no breaks, which is nested in eqs. (1)-(2) and serves as a benchmark. If the estimated sign of the parameter ς is positive, then inflation uncertainty will raise the inflation rate, supporting the stand of Cukierman and Meltzer (1986); however, if its sign is negative, inflation uncertainty will decrease the inflation rate, supporting the claim of Holland (1995). In model M2 the impact of inflation on its uncertainty is determined by the sign of coefficient d. If the Friedman (1977) (see also Ball, 1992) hypothesis holds in this model $d > 0$. Therefore, also model M2 is nested in eqs. (1)-(2) with no breaks. Note that in model M2 there is a bidirectional feedback between the conditional mean and variance (in-mean and level effects). Model M3 is similar to the specification model M2, but it allows for the impact of regime changes in the intrinsic persistence (a regime shift in 1990Q4) and the in-mean parameter (a regime change in 1954Q4), accordingly $\phi_1 \neq 0$ and $\varsigma_1 (= \varsigma_2) \neq 0$. Finally, model M4 allows for a break in the constant of the conditional mean equation in 1990Q4 (i.e. $\varphi_1 \neq 0$) and regime changes in the conditional variance equation in the level parameter ($d_1 \neq 0$) again in 1990Q4, in addition to a break in the variance persistence in 1954Q4 (i.e. $\beta_1 (= \beta_2) \neq 0$).

[13] Karanasos and Schurer (2008) show that it is optimal to model the conditional standard deviation of inflation instead of the conditional variance. So far the relevant empirical literature has ignored this important characteristic of the inflation data. We will use the acronym (P)ARCH, since δ is not estimated but it is set equal to one.

Table 2. Estimated DAB-AR$(1;2)$-M model using U.S. inflation data

	M1	M2	M3	M4
	Panel A: Conditional Mean			
φ	0.0004 (0.000)	0.0004 (0.000)	0.002^{**}	0.001^{*} (0.0006)
φ_1	$-$	$-$	$-$	-0.001^{*} (0.0004)
ϕ	0.769^{***} (0.048)	0.749^{***} (0.044)	0.598^{***} (0.048)	0.699^{***} (0.051)
ϕ_1	$-$	$-$	0.318^{***} (0.035)	$-$
ς	0.354^{*} (0.200)	0.396^{**} (0.169)	0.549^{***} (0.213)	0.434^{**} (0.200)
$\varsigma_1 = \varsigma_2$	$-$	$-$	0.233^{***} (0.022)	$-$
	Panel B: Conditional Variance			
ω	0.0002 (0.0008)	0.0001 (0.0001)	0.0011^{***} (0.000)	0.002 (0.000)
α	0.155^{***} (0.044)	0.148^{***} (0.049)	0.173^{***} (0.050)	0.145^{***} (0.045)
β	0.846^{***} (0.038)	0.797^{***} (0.053)	0.795^{***} (0.044)	0.825^{***} (0.047)
$\beta_1 = \beta_2$	$-$	$-$	$-$	0.758^{***} (0.038)
d	$-$	0.016^{***} (0.009)	0.019^{***} (0.009)	0.031^{***} (0.010)
d_1				0.024^{***} (0.009)
R^2	0.57	0.58	0.59	0.57
	Panel C: Q-Statistics and Information Criteria			
Q-Statistics (4)	0.734 [0.848]	0.768 [0.943]	3.735 [0.712]	4.734 [0.692]
Akaike	-8.438	-8.392	-8.384	-8.320
Schwarz	-8.300	-8.292	-8.272	-8.196

Note: The table reports the estimated parameters of models M1-M4 and the related standard errors . ***, **, * indicate statistical significance at 1%, 5% and 10%, respectively. In Panel C the calculated values of the Ljung-Box Q statistic and the p-values are reported along with the information criterion. The parameters in the three periods are as follows, i.e., for ϕ: $\phi = \phi_3$ in the pre-1955Q1 period, ϕ_2 in the second period, and ϕ_1 in the post 1990Q4 period.

Looking at the estimated parameters two important results arise from Table 2. First, the estimated in-mean parameter, ς (and ς_1), is positive and statistically significant in all the estimated models. The sign of the estimated parameter confirms the Cukierman-Meltzer hypothesis that uncertainty about future inflation positively affects inflation. However, the estimated parameters for d (and d_1) are also positive and significantly different from zero thus supporting the Friedman (1977) hypothesis that an increase in inflation may induce more uncertainty about future inflation. This result is consistent with the one in Conrad and Karanasos (2015a). [14] The estimation results in model M2 also suggest that estimating the conditional mean and variance equation independently would obtain a biased estimate of the degree of persistence since from Proposition 3 it is clear that y_t and σ_t share the same autoregressive polynomial.

Coming now to Model 3 and Model 4, the former suggests a sharp decrease of intrinsic persistence in the 90s since the estimated parameter $\phi_1 < \phi$ and significantly different from zero.[15] Similarly,

[14] In the literature Kontonikas (2004), Daal et al. (2005), and Fountas and Karanasos (2007) also find evidence of the positive effect of inflation on inflation uncertainty.

[15] Note that attempts to estimate models with regime changes in ω and α returned estimated parameters that were not

according to model M3 the estimated coefficient $\varsigma_1 (= \varsigma_2) < \varsigma$, thus suggesting a smaller impact of inflation uncertainty after the 1954. Model M4 tells a similar story, since it also suggests a decrease in the persistence of the conditional variance in the post-1954 period and a regime change in the level parameter in the 1990s (i.e. $d > d_1$).

The empirical estimates in models M3 and M4 are consistent with the hypothesis by Nel Negro *et al.* (2020) of a flattening of the Phillips curve around the 1990s and it is in line with a general reading of Federal Reserve policy history. Before the 1990s, inflation persistence was relatively high, suggesting that the preference for inflation stability was weak relative to the goal of stabilizing output. A decline in persistence started with so called Volcker disinflation period suggesting that the Federal Reserve became substantially more concerned with inflation stabilization after the 1980s. Inflation and inflation persistence remained low during Greenspan's tenure, suggesting continued strong preferences for inflation stability and transparence of monetary policy (see McLeay and Tenreyro, 2019). Other factors in the economic environment may also have been at play, helping to drive inflation persistence lower after the 1990s. The results support the claim in Stock and Watson (2007) that inflation has been much less volatile in the last thirty years than it was in the 1970s or early 1980s.

Looking now at the specification tests in Panel C the reported values of the Ljung-Box Q statistic do not reject the null hypothesis that there is no autocorrelation up to the forth order, thus indicating the absence of serial correlation.

4.3 Forecasting Exercise

A rolling forecast experiment is implemented in order to compare the forecasting ability of the different model specifications in Table 2. That is, the out-of-sample predictive properties of the estimated models are investigated via a rolling forecast experiment.

The forecasting exercise is based on the fixed-rolling windows scheme by Rossi and Sekhposyan (2010): in a set of periods $\{1, ..., T\}$, we produce a number P of forecasts obtained by using estimates of a regression. Thus, there are P out-of-sample predictions to be evaluated, where the first out-of-sample prediction is based on a parameter estimated using data generated by the estimated model up to time R; the second prediction is based on a parameter estimated using data up to $R+1$, and the last prediction is based on a parameter estimated using data up to $R+P-1=T$, where $T^s = R+P+h-1 = T+h$ is the size of the available sample and $h = \{3, 8, 12\}$ being the pseudo-out-of-sample horizon. Our data spans from 1947Q2 to 2021Q3 so that $T = 297$ is the *in-sample* part; the *estimation* part spans from 1947Q3 to 1983Q4 (corresponding to $R = 147$) and the *evaluation* part from 1984Q1 to 2018Q3 (corresponding to $P = 138$). These last are the basis to compute the aggregate measures of point and density forecasts which allow to evaluate the best performing model.

significantly different from zero.

The rolling-window forecast scheme was preferred to a recursive (or expanding) scheme, since the latter is better able to handle the parameter switching in the inflation series. Moreover, it is consistent with the conventional view that more recent observations are more informative than those at the very beginning of a sample period. The out-of-sample forecast comparisons do not rely on a single criterion, for robustness we compare the results using point and density measures of accuracy.[16]

a) Point Forecasts Measures

Let $\hat{y}_t$ the estimated inflation series of each model j, the point predictive performances of model are investigated using three different measures: the mean forecast error (MFE), the symmetric mean absolute percentage error (sMAPE) and the median relative absolute error (mRAE). The three performance measures are calculated as follows:

$$\begin{aligned} MFE_{j,h} &= \frac{1}{T-h-T^s+1}\sum_{t=T^s}^{T-h}\left(y_{t+h}-\hat{y}^{(j)}_{t+h|t}\right), \\ sMAPE_{j,h} &= \frac{100|y_{t+h}-\hat{y}^{(j)}_{t+h}|}{0.5(y_{t+h}-\hat{y}^{(j)}_{t+h|t})}, \\ mRAE_{j,h} &= \frac{|y_{t+h}-\hat{y}^{(j)}_{t+h}|}{|y_{t+h}-\hat{y}^{(1)}_{t+h}|}, \quad \text{with (1) indexing the benchmark model;}. \end{aligned}$$

b) Density Forecast Measures

The literature on the aggregation of density forecasts focuses on the so-called scoring rules (see, for example, Geweke and Amisano, 2011). These are functions that enable the forecaster to aggregate the set of conditional predictive densities. As for point forecasting, the out-of-sample forecast comparisons are based on four different scoring rules are used for aggregating the $T-T^s-h+1$ predictive densities produced by the same forecasting exercise:

The logarithmic score (LogS):

$$LogS_{j,h} = \frac{1}{T-h-T^s+1}\sum_{t=T^s}^{T-h}\log f^{(j)}_{t+h|t}, \tag{21}$$

which corresponds to a Kullback-Liebler distance from the true density; models with higher LogS are preferred.

The quadratic score, somewhat the equivalent of the MSFE in point forecasting, is defined as:

$$QRS_{j,h} = \frac{1}{T-h-T^s+1}\sum_{t=T^s}^{T-h}\sum_{k=1}^{K}\left(f^{(j)}_{t+h|t}-d_{kt}\right)^2,$$

[16] The literature on point forecasting and the evaluation of individual density forecasts is well established; see for example Corradi and Swanson (2006).

where $d_{kt} = 1$ if $k = t$ and 0 otherwise; models with lower QSR are preferred.

The (aggregate) continuous-ranked probability score (CRPS), equivalent to the sMAPE, is given by:

$$CRPS_{j,h} = \frac{1}{T - h - T^s + 1} \times \sum_{t=T^s}^{T-h} \left(\left| f_{t+h} - f_{t+h|t}^{(j)} \right| - 0.5 \left| f_{t+h} - f_{t+h|t}^{'} \right| \right),$$

where f and f' are independent random draws from the predictive density and $f_{t+h|t}$ is the observed; models with lower CRPS are preferred.

Finally, the quantile score (qS), which can be obtained if $f_{t+h|t}^{(j)}$ is replaced with a predictive α-level quantile $q_{t+h|t}^{\alpha}$ in eq. (21) (and the logarithmic function is removed); this score is used in risk analysis because it provides information about deviations from the true tail of the distribution. The lower the score, the better the forecasts are.

Table 3 reports the results of the h-step-ahead forecasts for the forecast period $h = \{3, 8, 12\}$. In Panel A the point forecast measures are reported, whereas density forecast performance measure are reported in Panel B. In columns 1 and 2 the forecasting horizon and the forecast error measures are reported, respectively, whereas in columns 3-6 the forecasting results for each model are reported.

Table 3. Forecasting inflation: point and density predictive performances for models M1-M4.

Forecast Horizon	Forecast Error Measure	M1	M2	M3	M4
PANEL A: Point Forecasts					
3	MFE	0.0020	0.0009	0.0005	0.0005
8		0.0026	0.0016	0.0012	0.0011
12		0.0033	0.0027	0.0018	0.0018
3	sMAE	0.0045	0.0021	0.0018	0.0020
6		0.0052	0.0024	0.0022	0.0025
12		0.0057	0.0025	0.0024	0.0027
3	mRAE	1.000	0.9450	0.9539	0.9511
8		1.000	0.9551	0.9610	0.9588
12		1.000	0.9555	0.9700	0.9626
PANEL B:Density Forecast					
3	LogS	0.1459	0.1452	0.1430	0.1435
8		0.1500	0.1427	0.1422	0.1424
12		0.1588	0.1442	0.1440	0.1429
3	QRS	0.6940	0.6803	0.6774	0.6771
8		0.7020	0.699	0.6997	0.6835
12		0.7122	0.5063	0.7332	0.6956
3	CRPS	0.8900	0.8914	0.8930	0.8913
8		0.9063	0.9060	0.9035	0.8925
12		0.9134	0.9181	0.9093	0.9065
3	qS	0.0240	0.0235	0.0235	0.0237
8		0.0262	0.0260	0.0259	0.0264
12		0.0265	0.0263	0.0260	0.0271

The table compares the four different models in their out-of-sample forecasts. In Panel A the point forecast measures are *i*) the mean forecast error (MFE), *ii*) the symmetric mean absolute percentage error (sMAPE), and *iii*) the median relative absolute error (mRAE). In Panel B the density forecast measures are: *i*) the logarithmic score (LogS), *ii*) the quadratic score (QSR), *iii*) the continuous-ranked probability score (CRPS), and *iv*) the quantile score (qS). The forecast horizon is 3, 8, and 12 quarters.

From Panel A of Table 3 it is clear that according to the point performance measures, models that allow for structural breaks perform better than their counterpart in all cases but one. The density performance measures reported in Panel B are in line with the results of the point loss functions, thus confirming that models that allow for time varying parameters enjoy better forecast performance, especially for medium and long run horizons.

Given the ranking obtained in the previous section, we now investigate whether the predictive densities in Table 3 are significantly different from each others. With this target in mind we investigate weather the specifications M3 and M4 perform better than M1 and M2 using three tests: the DM test (Diebold and

Mariano, 1995), the GW test (Giacomini and White, 2006) and the AG test (Amisano and Giacomini, 2007) for equal predictive ability for pairs of forecast.

In brief, the DM test has under the null hypothesis that the forecast errors coming from the two forecasts bring about the same loss. Under the assumption that the loss differential is a covariance stationary series, the loss differential converges asymptotically to a normal distribution. A possible drawback of this test is that the procedure has been designed for non nested models and may be not accurate in our case. Accordingly, we also consider the GW and AG tests inference procedures that are valid under more general conditions. In addition, the forecasts can be based on nested or nonnested models.

The GW test can be seen as a generalization of the DM test and it measures the conditional predictive ability rather than the unconditional predictive ability. Like the DM procedure, the test measures the statistical significance of the differences of two models forecasts. Under the null hypothesis the test is asymptotically distributed as a χ^2-distribution. Unlike the previous tests, the AG test is an inference procedure useful for comparing the out-of-sample accuracy of competing density forecasts since it allows to test if the predictive densities of competing models are significantly different from each others. The evaluation is based on scoring rules, which are loss functions defined over the density forecast and the realizations of the variable (see, for details, Amisano and Giacomini, 2007).

Table 4 reports the estimated p-values of the tests. The first column reports the forecast horizon, whereas columns 2-6 report the p-values for different null hypotheses. In column two, under the null hypothesis the conditional predictive ability of the loss differential of model M1 is higher or equal than that of model M2. Hence, rejecting the null means that the forecasts of model M2 are significantly more accurate than those of model M1. Likewise, in column three and four, under the null hypothesis the conditional predictive ability of the loss differential of model M1 is higher or equal to the conditional predictive ability of model M3 and model M4, respectively. The goal of this exercise is to investigate if the benchmark model M1 outperforms the other models. Finally, in the last two columns we investigate if models M2 and M4 outperform model M2.

Table 4. Predictive ability tests for different models in quarterly inflation data.

Forecast Horizon	M2 vs M1	M3 vs M1	M4 vs M1	M3 vs M2	M4 vs M2
		Diebold-Mariano			
3	0.035	0.052	0.004	0.041	0.039
8	0.053	0.048	0.031	0.054	0.050
12	0.075	0.048	0.045	0.058	0.068
		Giacomini-White			
3	0.045	0.003	0.009	0.044	0.046
8	0.074	0.046	0.030	0.067	0.060
12	0.099	0.033	0.033	0.099	0.066
		Amisano-Giacomini			
LogS					
3	0.041	<0.000	<0.000	0.013	0.028
8	0.021	<0.000	<0.000	0.041	0.040
12	0.010	<0.000	<0.000	0.048	0.052
QRS					
3	0.034	0.007	0.006	0.032	0.065
8	0.045	0.012	0.034	0.048	0.083
12	0.091	0.030	0.057	0.055	0.094
CRPS					
3	0.153	0.142	0.113	0.092	0.084
8	0.124	0.104	0.091	0.135	0.099
12	0.092	0.099	0.053	0.136	0.125
qS					
3	0.013	0.020	0.030	0.056	0.056
8	0.022	0.043	0.041	0.071	0.068
12	0.019	0.044	0.041	0.078	0.076

NOTE: The table reports the results of equal predictive ability tests for models M1-M4 on quarterly U.S. inflation data for the testing period from 1947:Q1 to 2021:Q3. All of the tests consider density forecasts generated according to the models estimated in Table 2.

Looking at the estimated p-values in Table 4 it is clear the all three tests reject the null hypothesis of equal conditional predictive ability of the benchmark model M1, thus we can conclude that when it comes to forecasting inflation it is important taking into consideration the impact of inflation on its uncertainty. Looking now at the last two columns of Table 4 it is clear that models M3 and M4 outperform model M2. Overall, model M3 has the smallest forecast errors and is preferred in almost all forecast horizons according to the other point forecast measures. This prominence is broken by model M4 if considering density forecasts.

5 Forecasting Evaluation of Alternative Models

In the previous section we have shown that the DAB-AR-ML model captures the relation between inflation and its uncertainty well and it has good predictive power. However, a useful forecasting exercise requires an evaluation against competitive models. Accordingly, in this section we investigate the forecasting properties of the DAB-AR-ML specification by comparing the forecasting performance of the suggested model to alternative linear and nonlinear specifications.

In the related literature the random walk (RW) has often been used in empirical studies to forecast inflation as a benchmark (see for example Atkeson and Ohanian, 2001, Fisher *et al.* 2009). In the literature autoregressive models are also found useful to forecast inflations in the United States; see for example Giordani (2003), and Orphanides and Van Norden (2005). Autoregressive processes have been used in the context of testing for the long-run effect of a shock to inflation, mainly using the LAR and SAR measures of persistence. For this reason, in this study the AR(p) is used as a benchmark model in the forecasting competition.[17] Coming to nonlinear specifications, modelling inflation using STAR-type models has become increasingly popular in recent years as these models allow for endogenous regime switching mechanism in the inflation persistence. Most empirical works accommodate the departure from linearity of the price change series by using a transition or switching mechanism that captures the fact that the inflation persistence behaves differently according to the state of the economic cycle (see for example Rossi and Sekhposyan, 2010; Hubrich and Skudelny, 2016). Therefore, the smooth transition autoregressive model (STAR) (see, for example, Teräsvirta *et al.*, 2005) is considered in this paper. Finally, we consider an autoregressive Markov regime switching GARCH-M model that is similar to the specification in eqs. (1) and (2), but it also allows for endogenous switching of the parameters.

As far as the model estimation is concerned, following Atkeson and Ohanian (2001), the RW model predicts that inflation over the next three quarters is expected to be equal to inflation over the previous three quarters. As for the autoregressive model the maximal lag order of the AR(4) model was chosen by using the Bayesian information criterion and the Portmanteau test for serial correlation.

Coming to the STAR model, the test for linearity suggested by Luukkonen *et al.* (1988) rejected the null hypothesis of linearity in the time series under consideration.[18] Following the model selection procedure in Teräsvirta *et al.* (2005) it was found that the STAR model with a logistic transition function (LSTAR model) was the best model specification. Accordingly the following model was estimated

$$y_t = \left[\varphi + \sum_{i=1}^{4} \phi_i y_{t-i}\right] + \left[\rho_0 + \sum_{i=1}^{4} \rho_i y_{t-i}\right] \left[1 + \exp\{-\lambda\}\left(y_{t-d} - g\right)\right]^{-1} + \varepsilon_t, \tag{22}$$

where $\lambda > 0$. In eq. (22) inflation evolves with a smooth transition between low and high persistence regimes that depends on the sign and magnitude of past realizations of price growth rates. The parameter λ denotes the speed of transition between regimes and g measures the threshold between the two regimes.

Considering now the regime switching model (MS-AGARCH-M) in this specification inflation persistence switches between high and low regimes according to an autoregressive Markov switching model. Following Chang (2012) (see also Ardia, 2008; Marcucci, 2005; Chan, 2012) we consider the following model

[17] Note that in the literature ARIMA processes have also been used to forecast inflation. The forecasting properties of this model in case of structural breaks are investigated in Canepa *et al.* (2019).

[18] The linearity test is not reported but it is available upon request.

$$y_t = \mu(s_{1t}) + \phi(s_{1t}) y_{t-1} + \varsigma(s_{1t}) \sigma(s_{2t}) + \varepsilon_t, \quad (23)$$

$$\sigma^2(s_{2t}) = \omega(s_{2t}) + \alpha(s_{2t}) \varepsilon_{t-1}^2 + \rho(s_{2t}) I_{(\varepsilon_{t-1}>0)} \varepsilon_{t-1}^2 + \beta(s_{2t}) \sigma_{t-1}^2. \quad (24)$$

Parameters shown in eq. (23) are influenced by state variable s_{1t}. $\mu(s_{1t})$ and $\phi(s_{1t})$ refer to the intercept term and autoregressive term, respectively. As before the parameter $\varsigma(s_{1t})$ represents the in-mean term and it reflects the regime-dependent effect of inflation uncertainty on inflation. Eq. (24) shows the regime-dependent variance equation. The conditional variance is an asymmetric GARCH specification, and it is related to the state variable s_{2t} but not to the state variable s_{1t}. $I_{(\varepsilon_{t-1}>0)}$ is an indicator variable with value 1, if $\varepsilon_{t-1} > 0$, and with value 0, if $\varepsilon_{t-1} \leq 0$. According to this model the impact of inflation on its uncertainty is replaced by the asymmetry term $\rho(s_{2t})$.[19] Finally, the state variable s_{1t} is assumed to have two different values. When its value is equal to 1, the economic system belongs to the increasing inflation pressure state (the inflation increases from the bottom to the peak). On the contrary, when its value is equal to 2, the economic system belongs to the decreasing inflation pressure state (see, for more details, Chang, 2012). The model has been estimated by QML under the assumption the error term follows a Student-t distribution.

The estimation of the parameters for the four different models are reported in Appendix 2, whereas in Table 5 the results of the forecasting exercise are given. In columns one and two the forecasting horizons and the forecast error measures are reported, whereas in columns 3-6 the forecasting results for each model are reported. Note that in Table 5 only the performance of models M3 and M4 are reported since according to the results in Table 4 these models are the preferred model specifications.[20]

[19] Note that unlike the DAB-AR-ML model the specification adopted by Chang (2012) does not allow to simultaneously tests both the impact of inflation on its uncertainty and the in-mean effect. Rather than focusing on the level effect, the model in Chang (2012) captures the impact of asymmetries to the conditional variance. However, in this paper we are more interested on the impact of inflation on its uncertainty. Therefore, we follow Fountas and Karanasos (2007) and introduce the lagged y_t directly into the conditional variance equation.

[20] Note that similar results for models M1 and M2 are available from the authors upon request.

Table 5. Forecasting performance measures for alternative estimated models.

Forecast Horizon	Forecast Error Measure	RW	AR(4)	LSTAR	MS-AGARCH-M
PANEL A: Point Forecasts					
3	MFE	0.0150	0.0155	0.0144	0.0075
8		0.0208	0.0189	0.0125	0.0102
12		0.0227	0.020	0.0179	0.0109
3	sMAE	0.0067	0.0028	0.0019	0.0022
8		0.0089	0.0046	0.0031	0.0048
12		0.0102	0.0050	0.0031	0.0050
3	mRAE	1.4553	1.2445	1.0975	1.0556
8		1.4885	1.2890	1.1355	1.0592
12		1.4896	1.2867	1.1367	1.0600
PANEL B:Density Forecast					
3	LogS	0.2446	0.1836	0.1764	0.1670
8		0.3055	0.1996	0.1963	0.1773
12		0.3055	0.2004	0.1970	0.1774
3	QRS	0.8944	0.7853	0.6998	0.6920
8		0.9935	0.8934	0.7100	0.7036
12		1.1306	0.9631	0.7250	0.7110
3	CRPS	1.0068	0.9133	0.9066	0.8958
8		1.0466	0.9268	0.9765	0.9230
12		1.3561	0.9278	0.9769	0.9295
3	qS	0.0350	0.0256	0.0254	0.0256
8		0.0500	0.0288	0.0273	0.0274
12		0.0588	0.0277	0.0275	0.0274

The table compares RW, AR(4), LSTAR, MS-AGARCH-M in their out-of-sample forecasts. In Panel A the point forecast measures are *i*) the mean forecast error (MFE), *ii*) the symmetric mean absolute percentage error (sMAPE), and *iii*) the median relative absolute error (mRAE). In Panel B the density forecast measures are: *i*) the logarithmic score (LogS), *ii*) the quadratic score (QSR), *iii*) the continuous-ranked probability score (CRPS), and *iv*) the quantile score (qS). The forecast horizon is 3, 8, and 12 quarters.

Looking at the results from Table 5 it appears that the random walk model has the worst performance in forecasting the inflation process. The AR(4) certainly bits the RW model in term of forecasting performance. Looking now at the models that allow for changes in the inflation persistence the LSTAR model appears to be the worst specification in term of forecasting performance, whereas the MS-AGARCH-M which allows for two states inflation uncertainty has better performance.[21] Overall, the DAB-AR-ML

[21] The choice of length of the "estimation" and "evaluation" parts of the samples allows us to avoid a bias in favor of the models M3 and M4 this is because the hight oscillations of the series in the first half of the sample (corresponding to the estimation part) could in principle, favor nonlinear models such as the LSTAR specification rather than GARCH families. The second half of the sample (corresponding to the evaluation part) has still some nonlinear features, but the change in the variance of the sample after mid-80s seems more compatible with the linear families of models. This may induce the reader to think that the model M3 may be spurred. We remark, however, that both the sub-samples are characterized by

model bits their competitors for most forecasting horizons (see also the last two columns of Tables 3 and 5).

Tables 6 and 7 report the estimated p-values of the DM, GW and the GA tests. The first column reports the forecast horizon, whereas columns 2-5 report the p-values for different null hypotheses. In column two under the null hypothesis the conditional predictive ability of the loss differential of model M3 (M4) is higher or equal to that of the RW model. Likewise, in columns three, four and five under the null hypothesis the conditional predictive ability of the loss differential of model M3 (M4) is higher or equal to the conditional predictive ability of the AR(4), the LSTAR and the MS-AGARCH-M model, respectively.

Table 6. Predictive ability tests for different models in quarterly inflation data.

Forecast Horizon	M3 vs RW	M3 vs AR(4)	M3 vs LSTAR	M3 vs MS-AGARCH-M
		Diebold-Mariano		
3	<0.000	0.004	0.012	1.000
8	<0.000	0.009	0.038	1.000
12	<0.000	0.010	0.040	1.000
		Giacomini-White		
3	<0.000	0.011	0.015	0.774
8	<0.000	0.010	0.031	0.889
12	<0.000	0.013	0.037	0.941
		Amisano-Giacomini		
LogS				
3	<0.000	<0.000	0.024	0.883
8	<0.000	<0.000	0.051	0.897
12	<0.000	<0.000	0.066	0.990
QRS				
3	<0.000	0.006	0.054	1.000
8	<0.000	0.034	0.062	1.000
12	<0.000	0.057	0.092	1.000
CRPS				
3	0.004	0.013	0.061	0.988
8	0.009	0.021	0.072	1.000
12	0.022	0.024	0.090	1.000
qS				
3	<0.000	<0.000	0.070	0.899
8	<0.000	<0.000	0.093	0.950
12	<0.000	0.002	0.121	1.000

NOTE: The table reports the results of equal predictive ability tests for models M4, RW,AR(4), MS-AGARCH-M on quarterly inflation data for the testing period from 1947:Q1 to 2021:Q3. The tests are based on the results of Tables 2,4,5 and the estimated parameters in Appendix 2.

a change in conditional variance, so that the forecasting exercises is effectively fair.

Table 7. Predictive ability tests for different models in quarterly inflation data.

Forecast Horizon	M4 vs RW	M4 vs AR(4)	M4 vs LSTAR	M4 vs MS-AGARCH-M
		Diebold-Mariano		
3	<0.000	0.039	0.015	0.960
8	<0.000	0.068	0.024	0.985
12	<0.000	0.083	0.036	0.987
		Giacomini-White		
3	<0.000	0.054	0.016	0.800
8	<0.000	0.071	0.021	0.861
12	<0.000	0.072	0.021	0.866
		Amisano-Giacomini		
LogS				
3	<0.000	<0.093	0.003	0.828
8	<0.000	<0.021	0.014	0.879
12	<0.000	<0.000	0.020	0.901
QRS				
3	<0.000	0.030	0.014	1.000
8	<0.000	0.056	0.024	1.000
12	<0.000	0.004	0.048	1.000
CRPS				
3	<0.000	0.080	0.041	0.924
8	<0.000	0.029	0.055	0.977
12	<0.000	0.005	0.059	0.995
qS				
3	<0.000	0.024	0.018	0.830
8	<0.000	0.046	0.033	0.849
12	<0.000	0.060	0.041	0.888

NOTE: The table reports the results of equal predictive ability tests for models M4, RW,AR(4), MS-GARCH on quarterly inflation data for the testing period from 1947:Q1 to 2021:Q3. The tests are based on the results of Tables 2,4,5 and the estimated parameters in Appendix 2.

Coming to the results of the battery of predictive ability tests, overall, the null of equal predictive ability of model M3 with respect to the other models is rarely accepted. The exceptions are constituted almost exclusively by MS-AGARCH-M (which is similar to M3 and M4 by construction). These results are in line with Pettenuzzo and Timmerman (2017) where it is found that accounting for time varying parameters in models that account for stochastic volatility lead to more accurate forecasts.

6 Estimating Persistence

In Table 2 we have presented two variations of the DAB-AR-ML time varying model that seem to capture well the characteristic features of the inflation process. We now employ the AR-(P)GARCH-ML model to compute the first-order persistence presented in Section 3.3, which not only distinguishes between changes in the dynamics of inflation and its volatility (and their persistence), but also allows for bidirectional feedback between inflation and its volatility. Although according to the information criteria the best performing model is model M4, it is of interest to calculate the persistence measure for the four model specifications in Table 2 as they capture four different scenarios for the inflation process.

Table 8 presents the time invariant (within each period) first-order measures of persistence for the four models under consideration. In the top part of Table 8, the first four columns report the LAR, the next four columns the $1/(1-SAR)$, whereas the first four columns in the bottom part report the mean inflation and the last four columns the sum of the Wold coefficients (or Green functions) (see eq. (18)).

Model M4, which is the preferred one, generates higher persistence than model M3 in all three periods. Interestingly, both models show high persistence in the pre-1955 period. The persistence decreases in the period 1955-1990 by 81% according to model M4 and 61% for model M3. It falls further in the post-1990 period by 11% and 45% for models M4 and M3, respectively.

Table 8. First-order persistence for each of the three periods and four models.

	LAR				$1/(1-SAR)$			
	M1	M2	M3	M4	M1	M2	M3	M4
1947Q1-1954Q4	0.970	0.946	0.961	0.987	53.323	29.866	31.933	106.310
1955Q1-1990Q4	0.970	0.946	0.945	0.928	53.323	29.866	12.260	20.577
1991Q1-2021Q3	0.970	0.946	0.940	0.918	53.323	29.866	6.779	18.313
	$\mathbb{E}(y_t)$				SGF			
1947Q1-1954Q4	1.184	0.491	4.474	21.083	12.161	9.583	9.820	27.792
1955Q1-1990Q4	1.184	0.491	1.735	4.048	12.161	9.583	4.769	7.703
1991Q1-2021Q3	1.184	0.491	0.946	3.602	12.161	9.583	2.601	6.856

Note: For each period, $n = 1, 2, 3,$ we use the expressions in eq. (18) to calculate the (within each period time invariant). First-order persistence for the four models. The four measures are: The largest AR root (LAR), $1/(1-SAR)$ where SAR is the sum of the AR coefficients, the mean inflation ($\mathbb{E}(y_t)$), and the sum of the Green functions (SGF).

For comparison Table 8 presents the time invariant first-order persistence for models M1 and M2. They overestimate first-order persistence (measured by either $1/(1-SAR)$ or SGF) in the second and third period, whereas they underestimate it in the pre-1955 period. Model M1 and M2 also underestimate the mean for all three periods.

Therefore our findings regarding the first-order persistence are in line with the findings of Benati (2008) (see also Angeloni et al., 2003). In sum our main conclusion is that for the two chosen time varying specifications (models M3 and M4) the first-order measures of persistence, do not remain unchanged throughout the whole period 1947-2021 but decrease after 1994 and fall further in the post-1990 period. These results are in line with Brainard and Perry (2000), Taylor (2000), and Kim *et al.* (2004), who also found evidence that U.S. inflation persistence post 1990th has been substantially lower than during the previous two decades.

7 Conclusions

In recent years economists have placed significant increasing emphasis on investigating structural shifts in the dynamics of the inflation process in the United States. A number of detailed and rigorous empirical studies regarding changes in inflation persistence have, however, reached diverging conclusions. Several studies find evidence of little or no change of inflation persistence over the past four decades, whereas others conclude that there has been a pronounced decline over the same period.

In this paper we have attempted to reconcile different strands of the literature by showing that seemingly conflicting results regarding changes in inflation persistence actually constitute two sides of the same problem. Economic theory suggests that changes in the level and the second conditional moment of inflation process should be interrelated. However, in the related literature change of persistence in the level and in the conditional variance are usually analyzed independently. In this paper we show that these changes are interlinked. We then proceeded by using the general solution of a DAB AR-(PGARCH) M model to compute time varying persistence measures that are able to take into account the presence of breaks both on the conditional mean and variance. Finally, comparing a number of competitive model specifications we show that models that allow for time varying persistence have better forecasting properties with respect to specifications that only allow for regime changes in the autoregressive parameters.

APPENDIX 1

In this appendix we present the proofs of Theorems 1 and **??**. We will prove Theorem 1 by induction with respect to k.

Proof. (**of Theorem** 1). Clearly, it holds for $k+r=1$: In eq. (7) setting $k=1$, and thus $r=k_1=k_2=0$ we obtain eq. (6) since $\mathbf{E}(\mathbf{y}_\tau\,|\mathcal{F}_{\tau-1}) = \boldsymbol{\varphi} + \mathbf{\Phi}\mathbf{y}_{\tau-1} + \mathbf{Z}\boldsymbol{\varepsilon}_{\tau-1}$ and $\mathbf{FE}(\mathbf{y}_\tau\,|\mathcal{F}_{\tau-k}) = \mathbf{J}\boldsymbol{\varepsilon}_\tau$.
Next if we assume that it holds for k, then it will suffice to prove that it also holds for $k+1$. First, rewrite eq. (6) as of time $\boldsymbol{\tau}-k$:

$$\mathbf{y}_{\tau-k} = \boldsymbol{\varphi}(\tau-k) + \mathbf{\Phi}(\tau-k)\mathbf{y}_{\tau-(k+1)} + \mathbf{J}\boldsymbol{\varepsilon}_{\tau-k} + \mathbf{Z}(\tau-k)\boldsymbol{\varepsilon}_{\tau-(k+1)}.$$

Substituting the above equation into eq. (7) using straightforward algebra shows that

$$\mathbf{y}_{\tau,k+1} = \mathbf{E}(\mathbf{y}_\tau\,|\mathcal{F}_{\tau-(k+1)}) + \mathbf{FE}(\mathbf{y}_\tau\,|\mathcal{F}_{\tau-(k+1)})$$

as claimed. ■

APPENDIX 2

Parameter estimates for alternative models with standard errors:

AR(4) Model:

$$y_t = \underset{(0.0004)}{0.0012^{***}} + \underset{(0.058)}{0.619^{***}}y_{t-1} + \underset{(0.067)}{0.079}y_{t-2} + \underset{(0.060)}{0.123^{***}}y_{t-3} - \underset{(0.060)}{0.115^{**}}y_{t-4},$$

$R^2 = 0.61$, Akaike criterion: -8.099, Schwarz criterion: -8.037.

LSTAR Model:

$$\begin{aligned} y_t &= \underset{(0.0006)}{0.0025^{***}} + \underset{(0.113)}{0.420^{***}} y_{t-1} + \underset{(0.133)}{0.322^{**}} y_{t-2} - \underset{(0.140)}{0.158} y_{t-3} - \underset{(0.134)}{0.158} y_{t-4} \\ &+ \underset{(0.0038)}{0.0038} + \underset{(0.190)}{0.484^{**}} y_{t-1} - \underset{(0.221)}{0.647^{***}} y_{t-2} - \underset{(0.222)}{0.796^{***}} y_{t-3} \\ &+ \underset{(0.194)}{0.379^{**}} y_{t-4} \times \left[1 - \exp\left(-\underset{(245.22)}{405.73^{*}}\left(y_{t-4} - \underset{(0.001)}{0.010^{***}}\right)\right)\right]^{-1}, \end{aligned}$$

$R^2 = 0.666$, Akaike criterion: -8.217, Schwarz criterion: -8.156.

MS-AGARCH-M Model:

$$\begin{aligned} y_t &= \underset{(0.000)}{0.0004^{*}_{s1}} + \underset{(0.061)}{0.665^{***}} y_{t-1,s_1} + \underset{(0.142)}{0.314^{*}} \sigma_{t,s_1}, \\ & \underset{(0.0002)}{0.0003_{s2}} + \underset{(0.141)}{0.579^{***}} y_{t-1,s_2} - \underset{(0.012)}{0.156^{***}} \sigma_{t,s_2}, \\ \sigma_t^2 &= \underset{(0.001)}{0.0005^{*}_{s1}} + \underset{(0.077)}{0.199^{***}} \varepsilon_{t-1,s_1} - \underset{(0.051)}{0.243^{***}} + \underset{(0.011)}{0.429^{**}} \sigma^2_{t-1,s_1}, \\ & \underset{(0.000)}{0.001^{*}_{s2}} + \underset{(0.015)}{0.014^{***}} \varepsilon_{t-1,s_2} + \underset{(0.049)}{0.063} + \underset{(0.153)}{0.625^{**}} \sigma^2_{t-1,s_2}, \end{aligned}$$

The estimated transition probability matrices for s_{1t} and s_{2t} are:

$$Pr_1 = \begin{bmatrix} P(s_{1,t} = 1 | s_{1,t-1} = 1) & P(s_{1,t} = 1 | s_{1,t-1} = 2) \\ P(s_{1,t} = 2 | s_{1,t-1} = 1) & P(s_{1,t} = 2 | s_{1,t-1} = 2) \end{bmatrix} = \begin{bmatrix} \frac{\exp(0.9906)}{1+\exp(0.9906)} & \frac{1}{1+\exp(0.0094)} \\ \frac{1}{1+\exp(0.9906)} & \frac{0.0094}{1+\exp(0.0094)} \end{bmatrix},$$

and

$$Pr_2 = \begin{bmatrix} P(s_{2,t} = 1 | s_{2,t-1} = 1) & P(s_{2,t} = 1 | s_{2,t-1} = 2) \\ P(s_{2,t} = 2 | s_{2,t-1} = 1) & P(s_{2,t} = 2 | s_{2,t-1} = 2) \end{bmatrix} = \begin{bmatrix} \frac{\exp(0.2327)}{1+\exp(0.2327)} & \frac{1}{1+\exp(0.7673)} \\ \frac{1}{1+\exp(0.2327)} & \frac{0.7673}{1+\exp(0.7673)} \end{bmatrix}.$$

where $Pr(s_{1,t} = 2|_{s1,t-1} = 1)$ denotes the probability that state variable s_1 switches from state 1 to state 2. Similarly, $Pr(s_{2,t} = 1|s_{2,t-1} = 2)$ denotes the probability that state variable s_2 switches from state 2 to state 1. Akaike criterion: -8.318, Schwarz criterion: -8.221.

APPENDIX 3: Second Moment Structure

3.1 Time Invariant Case

The formulation in eq. (6) in the main body of the paper allows us to highlights the properties of the second moment structure of the DAB-BVARMA representation. First we will introduce some further notation.

Let $\mathbf{X}^{\otimes 2} = \mathbf{X} \otimes \mathbf{X}$ where $\otimes$ is the Kronecker product. Let also $vec(\mathbf{X})$ be a vector in which the columns of matrix $\mathbf{X}$ are stacked one underneath the other. Set $\mathbf{s}_\tau = vec(\mathbf{\Sigma}_\tau)$, see eq. (4) in the main body of the paper. In addition, let $\mathbf{\Gamma}_\tau$ denote the zero order bidimensional time varying covariance matrix of $\{\mathbf{y}_\tau\}$ and $\boldsymbol{\gamma}_\tau = vec(\mathbf{\Gamma}_\tau)$, that is $\boldsymbol{\gamma}_\tau = (\mathbb{V}ar(y_\tau), \mathbb{C}ov(y_\tau, \sigma_\tau^\delta), \mathbb{C}ov(y_\tau, \sigma_\tau^\delta), \mathbb{V}ar(\sigma_\tau^\delta))'$. For $\tau \leq t - k_2$ (the time invariant case) we set $\boldsymbol{\gamma} = \boldsymbol{\gamma}_\tau$ and $\mathbf{s} = \mathbf{s}_\tau$.

Assumption 1 (Second-Order). We assume that $\lambda_{\max}(\mathbf{\Phi}^{\otimes 2}) < 1$.

Notation 4 *For ease of presentation we will use the notation* $\mathbf{G} = [g_{ij}]_{i,j=1,\ldots,4}$ *where*

$$\mathbf{G} = \mathbf{J}^{\otimes 2} + (\mathbf{I} - \mathbf{\Phi}^{\otimes \mathbf{2}})^{-1}(\mathbf{\Phi}\mathbf{J} + \mathbf{Z})^{\otimes 2}. \tag{25}$$

Theorem 2 *Consider the general model in eq. (6) in the main body of the paper. Then, under Assumption 1, for $\tau \leq t - k_2$ $\boldsymbol{\gamma}$ is given by*

$$\boldsymbol{\gamma} = \mathbf{G}\mathbf{s}. \tag{26}$$

Proof. Rewrite eq. (6) in the main body of the paper for $\tau \leq t - k_2$:

$$\mathbf{y}_\tau = \boldsymbol{\varphi} + \mathbf{\Phi}\mathbf{y}_{\tau-1} + \mathbf{J}\boldsymbol{\varepsilon}_\tau + \mathbf{Z}\boldsymbol{\varepsilon}_{\tau-1}. \tag{27}$$

In view of eq. (27) the autocovariance matrix, $\mathbf{\Gamma}$, under Assumption 1, is given by:

$$\mathbf{\Gamma} = \mathbf{\Phi}\mathbf{\Gamma}\mathbf{\Phi}' + \mathbf{J}\mathbf{\Sigma}\mathbf{J}' + \mathbf{Z}\mathbf{\Sigma}\mathbf{Z}' + \mathbf{Z}\mathbf{\Sigma}(\mathbf{\Phi}\mathbf{J})' + \mathbf{\Phi}\mathbf{J}\mathbf{\Sigma}\mathbf{Z}'.$$

Applying the vec operator on both sides of the above equation we obtain

$$\boldsymbol{\gamma} = \mathbf{\Phi}^{\otimes \mathbf{2}}\boldsymbol{\gamma} + (\mathbf{J}^{\otimes 2} + \mathbf{Z}^{\otimes 2} + \mathbf{\Phi}\mathbf{J} \otimes \mathbf{Z} + \mathbf{Z} \otimes \mathbf{\Phi}\mathbf{J})\mathbf{s},$$

which implies that

$$\begin{aligned}
\boldsymbol{\gamma} &= (\mathbf{I} - \mathbf{\Phi}^{\otimes \mathbf{2}})^{-1}(\mathbf{J}^{\otimes 2} + \mathbf{Z}^{\otimes 2} + \mathbf{\Phi}\mathbf{J} \otimes \mathbf{Z} + \mathbf{Z} \otimes \mathbf{\Phi}\mathbf{J})\mathbf{s} \\
&= [\mathbf{J}^{\otimes 2} + (\mathbf{I} - \mathbf{\Phi}^{\otimes \mathbf{2}})^{-1}(\mathbf{\Phi}\mathbf{J}^{\otimes 2} + \mathbf{Z}^{\otimes 2} + \mathbf{\Phi}\mathbf{J} \otimes \mathbf{Z} + \mathbf{Z} \otimes \mathbf{\Phi}\mathbf{J})]\mathbf{s} \\
&= [\mathbf{J}^{\otimes 2} + (\mathbf{I} - \mathbf{\Phi}^{\otimes \mathbf{2}})^{-1}(\mathbf{\Phi}\mathbf{J} + \mathbf{Z})^{\otimes 2}]\mathbf{s}.
\end{aligned}$$

Thus in light of notation 4 the vec of the autocovariance matrix $\mathbf{\Gamma}$ can be written as

$$\boldsymbol{\gamma} = \mathbf{G}\mathbf{s}.$$

This completes the proof of theorem 2. ■

The Second Moment

The second moment of the power transformed conditional variance, $\mu_2 = \mu_{2\tau}$ for $\tau \leq t - k_2$, can be obtained as follows. First, notice that the fourth element of $\boldsymbol{\gamma}$ is the time invariant unconditional powered variance, $\mathbb{V}ar(\sigma_\tau^\delta) = \mu_2 - \mu_1^2$. We recall that for $\tau \leq t - k_2$, $\mathbf{\Sigma} = \mathbf{\Sigma}_\tau$, is given by

$$\mathbf{\Sigma} = \begin{bmatrix} \mu_{2/\delta}E(e_\tau^2) & \mu_{1+1/\delta}\widetilde{\kappa} \\ \mu_{1+1/\delta}\widetilde{\kappa} & \mu_2\kappa \end{bmatrix}$$

(see eq. (4) in the main body of the paper)

Next we will consider two special cases.

Case 1 $\delta = 1$ *and* $\gamma \neq 0$. *For this case* $\mathbf{\Sigma}$ *becomes*

$$\mathbf{\Sigma} = \mu_2 \begin{bmatrix} 1 & \widetilde{\kappa} \\ \widetilde{\kappa} & \kappa \end{bmatrix}.$$

Next we use the vec operator to get

$$\mathbf{s} = \mu_2 (1 \ \widetilde{\kappa} \ \widetilde{\kappa} \ \kappa)'. \tag{28}$$

Case 2 $\delta = 2$ *and* $\gamma = 0$. *For this case* $\boldsymbol{\sigma}$ *becomes*

$$\mathbf{s} = (\mu_1 \ 0 \ 0 \ \mu_2 \kappa)'. \tag{29}$$

Employing the notation $g = g_{41} + (g_{42} + g_{43})\widetilde{\kappa} + g_{44}\kappa$ (the g_{ij}'s have been defined in eq. (25)) the next theorem follows.

Theorem 3 *Consider the general model in eq. (6) in the main body of the paper. Then, under Assumption 1, for* $\tau \leq t - k_2$, μ_2 *is given by*

$$\mu_2 = \begin{cases} \text{1st Case:} & \frac{\mu_1^2}{1-g} & \text{iff } \ g < 1, \\ \text{2nd Case:} & \frac{\mu_1(\mu_1 + g_{41})}{1 - g_{44}\kappa} & \text{iff } \ g_{44}\kappa < 1. \end{cases} \tag{30}$$

Proof. 1st Case. On account of eqs. (26) and (28), the fourth element of $\boldsymbol{\gamma}$, if and only if $g < 1$, is given by

$$\mu_2 - \mu_1^2 = g\mu_2 \Rightarrow \mu_2 = \frac{\mu_1^2}{1-g}.$$

2nd Case. In light of eqs. (26) and (29), the fourth element of $\boldsymbol{\gamma}$, if and only if $g_{44}\kappa < 1$, is given by

$$\mu_2 - \mu_1^2 = g_{41}\mu_1 + g_{44}\mu_2\kappa \Rightarrow \mu_2 = \frac{\mu_1(\mu_1 + g_{41})}{1 - g_{44}\kappa},$$

as required. ■

Proposition 4 *Consider the general model in eq. (6) in the main body of the paper. Then, when* $d = 0$, *that is there are no level effects, if and only if* $c^2 + \alpha^2\kappa < 1$, μ_2 *for* $\tau \leq t - k_2$ *is given by*

$$\mu_2 = \frac{(1+c)\omega^2}{(1-c)(1-c^2-\alpha^2\kappa)},$$

which is a standard result (see, e.g., Karanasos, 1999, He and Teräsvirta, 1999, and Karanasos and Kim, 2006).

Proof. In the absence of level effects, since the matrix $\mathbf{\Phi}$ is upper triangular, the $\mathbf{G}$ matrix is also upper triangular and its $(4,4)$ time invariant element is $g_{44} = \frac{\alpha^2}{1-c^2}$. Thus, the fourth element of $\boldsymbol{\gamma}$, which is the time invariant unconditional variance of σ_τ^δ, is given by

$$\mathbb{V}ar(\sigma_\tau^\delta) = \mu_2 - \mu_1^2 = \frac{\alpha^2\kappa}{1-c^2}\mu_2.$$

Using $\mu_1 = \frac{\omega}{1-c}$ we obtain (if and only if $c^2 + \alpha^2\kappa < 1$) by straightforward manipulation:

$$\mu_2 = \frac{(1+c)\omega^2}{(1-c)(1-c^2-\alpha^2\kappa)},$$

and the proof is completed. ■

In the next section we will show how the above results can be used to derive $\mathbf{\Gamma}_\tau$ when $\tau > t - k_1$, that is when we have variable coefficients.

A 3.2. Time Varying Case

Before proceeding further, some additional notation is required.

Set $t + r - k = t - k_2$. Thus $k = r + k_2$. Since $\tau = t + r$, we also have: $\tau - k = t - k_2$.

Notation 5 *i) Denote the time varying coefficient matrices in the forecast error expansion (see Theorem 1 in the main body of the paper), that is the Green Matrices, by* $\mathbf{G}(\tau, \tau - l)$*. In other words,* $\mathbf{G}(\tau, \tau - l)$ *are given by*

$$\mathbf{G}(\tau, \tau - l) = \begin{cases} \mathbf{J} & \text{for} \quad l = 0, \\ \mathbf{\Phi}_1^{\ell-1}(\mathbf{\Phi}_1\mathbf{J} + \mathbf{Z}_1) & \text{for} \quad l = 1, \dots, k_1 + r, \\ \mathbf{\Phi}_1^{k_1+r}\mathbf{\Phi}_2^{l-k_1-r-1}(\mathbf{\Phi}_2\mathbf{J} + \mathbf{Z}) & \text{for} \quad l = k_1 + r + 1, \dots, k_2 + r - 1, \end{cases}$$

ii) In the sequel, for convenience and ease of exposition, we will use the notation

$$\mathbf{V}(\tau, \tau - l) = \mathbf{G}^{\otimes 2}(\tau, \tau - l), \tag{31}$$

that is, $\mathbf{V}(\tau, \tau - l) = [v_{ij}(\tau, \tau - l)]_{i,j=1,\dots,4}$ *is a time varying square matrix of order* 4.

Notation 6 *i) Henceforward we will make use of the notation:* $\mathbf{\Phi}(\tau, \tau - k) = \mathbf{\Phi}_1^{k_1+r}\mathbf{\Phi}_2^{k_2-k_1}$
ii) For notational ease we will set

$$\mathbf{\Lambda}(\tau, \tau - k) = [\mathbf{\Phi}(\tau, \tau - k)\mathbf{\Phi}_2]^{\otimes 2}\mathbf{G} + [\mathbf{\Phi}(\tau, \tau - k)\mathbf{Z}]^{\otimes 2}. \tag{32}$$

In other words, $\mathbf{\Lambda}(\tau, \tau - k) = [\lambda_{ij}(\tau, \tau - k)]_{i,j=1,\dots,4}$ *is a time varying square matrix of order* 4.

In the following theorem we make use of the two notations above.

Theorem 4 $\boldsymbol{\gamma}_\tau$*, that is vec form of the covariance matrix of* $\{\mathbf{y}_\tau\}$*, for* $\tau = t + r$*, is given by*

$$\boldsymbol{\gamma}_\tau = \sum_{l=0}^{k_2+r-1} \mathbf{V}(\tau, \tau - l)\mathbf{s}_{\tau-l} + \mathbf{\Lambda}(\tau, \tau - k)\mathbf{s}, \tag{33}$$

where the summation term coincides with the vec form of the $\mathbf{V}ar[\mathbf{FE}(\mathbf{y}_\tau | \mathcal{F}_{\tau-k})]$*, the second term on the right-hand side of eq. (33) concur with the vec form of the* $\mathbf{V}ar[\mathbf{E}(\mathbf{y}_\tau | \mathcal{F}_{\tau-k})]$*, and* $\mathbf{s}$ *has been derived in subsection 7, see eqs. (28) and (29), and eq. (30) in theorem 3.*

Proof. Theorem 1 in the main body of the paper implies that

$$\mathbf{\Gamma}_\tau = \mathbf{V}ar[\mathbf{E}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}\,)] + \mathbf{V}ar[\mathbf{FE}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}\,)],$$

or equivalently, by applying the vec operator

$$\boldsymbol{\gamma}_\tau = Vec\{\mathbf{V}ar[\mathbf{E}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}\,)]\} + Vec\{\mathbf{V}ar[\mathbf{FE}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}\,)]\}. \tag{34}$$

It remains to be shown that the first(second) term on the right-hand side of the above equation is equal to the first(second) term on the right-hand side of eq. (33). Utilizing notation 5(i) the forecast error (see theorem 1 in the paper), for $\tau > t - k_1$, can be written as

$$\mathbf{FE}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}\,) = \mathbf{J}\boldsymbol{\varepsilon}_\tau + \sum_{\ell=1}^{k_2+r-1} \mathbf{G}(\tau, \tau-l)\boldsymbol{\varepsilon}_{\tau-\ell}.$$

In view of the above equation the covariance matrix of the forecast error is given by

$$\mathbf{V}ar[\mathbf{FE}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}\,)] = \sum_{l=0}^{k_2+r-1} \mathbf{G}(\tau, \tau-l)\mathbf{\Sigma}_{\tau-l}\mathbf{G}'(\tau, \tau-l).$$

By applying the vec operator we obtain (in view of notation 5(ii))

$$Vec\{\mathbf{V}ar[\mathbf{FE}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}\,)]\} = \sum_{l=0}^{k_2+r-1} \mathbf{V}(\tau, \tau-l)\mathbf{s}_{\tau-l}. \tag{35}$$

Next we will make use of the notation 6(i) to write the optimal linear predictor in theorem 1 in the main body of the paper, for $\tau > t - k_1$, as

$$\mathbf{E}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}\,) = \mathbf{\Phi}(\tau, \tau-k)(\mathbf{\Phi}\mathbf{y}_{\tau-k} + \mathbf{Z}\boldsymbol{\varepsilon}_{\tau-k}).$$

Accordingly, its covariance matrix is given by

$$\mathbf{V}ar[\mathbf{E}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}\,)] = \mathbf{\Phi}(\tau, \tau-k)(\mathbf{\Phi}_2\Gamma\mathbf{\Phi}_2' + \mathbf{Z}\mathbf{\Sigma}\mathbf{Z}')\mathbf{\Phi}'(\tau, \tau-k).$$

Next we use the vec operator to get

$$\begin{aligned}
Vec\{\mathbf{V}ar[\mathbf{E}(\mathbf{y}_\tau \,|\mathcal{F}_{\tau-k}\,)]\} &= [\mathbf{\Phi}(\tau, \tau-k)\mathbf{\Phi}_2]^{\otimes 2}\boldsymbol{\gamma} + [\mathbf{\Phi}(\tau, \tau-k)\mathbf{Z}]^{\otimes 2}\mathbf{s} \\
\text{(by virtue of eq. (26))} &= \left\{[\mathbf{\Phi}(\tau, \tau-k)\mathbf{\Phi}_2]^{\otimes 2}\mathbf{G} + [\mathbf{\Phi}(\tau, \tau-k)\mathbf{Z}]^{\otimes 2}\right\}\mathbf{s} \\
\text{(on account of eq. (32))} &= \mathbf{\Lambda}(\tau, \tau-k)\mathbf{s}.
\end{aligned} \tag{36}$$

By virtue of eqs. (34), (35) and (36) it follows that

$$\boldsymbol{\gamma}_\tau = \sum_{l=0}^{k_2+r-1} \mathbf{V}(\tau, \tau-l)\mathbf{s}_{\tau-l} + \mathbf{\Lambda}(\tau, \tau-k)\mathbf{s},$$

as required. ∎

The Time Varying Second Moment

Lastly, we need to obtain $\mathbf{s}_{\tau-l}$. We will consider the two cases 1 and 2. We recall that for the second case, that is the case with no asymmetries, where $\delta = 2$ and $\gamma(\tau) = 0$ for all τ, $\mathbf{s}_\tau$ becomes (see eq. (29))

$$\mathbf{s}_\tau = (\mu_{1\tau} \ 0 \ 0 \ \mu_{2\tau}\kappa)'. \tag{37}$$

To proceed we introduce the following notations.

Notation 7 *For convenience of presentation we will*
i) set $p = k_2 + r - 1$,
ii) use the notation

$$c(\tau) = \mu_{1\tau}^2 + \sum_{l=1}^{p} v_{41}(\tau, \tau - l)\mu_{1,\tau-l} + \lambda_{41}(\tau, \tau - k)\mu_1 + \lambda_{44}(\tau, \tau - k)\kappa\mu_2 \tag{38}$$

(suppressing the dependence of $c(\tau)$ *on* k*),*
iii) set

$$\phi_l(\tau) = \kappa v_{44}(\tau, \tau - l)$$

(we recall that $v_{ij}(\tau, \tau - l)$ *and* $\lambda_{ij}(\tau, \tau - k)$ *are given in notations 5(ii) and 6(ii), respectively; see eqs. (31) and (32)).*

Proposition 5 *The time varying second moment of* σ_τ^2*, that is* $\mu_{2\tau}$*, when* $\delta = 2$ *and* $\gamma(\tau) = 0$ *for all* τ*, obeys a time varying difference (TV-DE) equation of order* p *with initial values* μ_2 *as follows*

$$\mu_{2\tau} = c(\tau) + \sum_{l=1}^{p} \phi_l(\tau)\mu_{2,\tau-l}. \tag{39}$$

Proof. Observe that the fourth element of $\boldsymbol{\gamma}_\tau$ in eq. (33) is $\mu_{2\tau} - \mu_{1\tau}^2$ and reiterate that the elements in the fourth row of $\mathbf{V}(\tau, \tau - l)$ are denoted as $v_{4j}(\tau, \tau - l)$. Consequently, by virtue of eqs. (33) and (37) $\mu_{2\tau}$ is given by

$$\mu_{2\tau} = \mu_{1\tau}^2 + \sum_{l=1}^{k_2+r-1} v_{41}(\tau, \tau - l)\mu_{1,\tau-l} + \lambda_{41}(\tau, \tau - k)\mu_1 + \lambda_{44}(\tau, \tau - k)\kappa\mu_2 + \kappa \sum_{l=1}^{k_2+r-1} v_{44}(\tau, \tau - l)\mu_{2,\tau-l}. \tag{40}$$

In view of notation 7, $\mu_{2\tau}$ in eq. (40) can be written as

$$\mu_{2\tau} = c(\tau) + \sum_{l=1}^{p} \phi_l(\tau)\mu_{2,\tau-l},$$

as required. ■

The solution of the TV-DE in eq. (39) can be derived by applying the methodology recently introduced in Karanasos et al. (2022).

Set $s = t - k_2$. For every pair $(\tau, s) \in \mathbb{Z}^2$ such that $\tau - s \geq 1$ the principal matrix associated with the TV-DE in eq. (39), is defined by

$$\mathbf{\Phi}_{\tau,s} = \begin{bmatrix} \phi_1(s+1) & -1 & & & & & & \\ \phi_2(s+2) & \phi_1(s+2) & \ddots & & & & & \\ \vdots & \vdots & \ddots & \ddots & & & & \\ \phi_p(s+p) & \phi_{p-1}(s+p) & \ddots & \ddots & \ddots & & & \\ & \phi_p(s+p+1) & \ddots & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & \phi_p(\tau-1) & \phi_{p-1}(\tau-1) & \cdots & \phi_1(\tau-1) & -1 \\ & & & & \phi_p(\tau) & \cdots & \phi_2(\tau) & \phi_1(\tau) \end{bmatrix}, \quad (41)$$

here and in what follows empty spaces in a matrix have to be replaced by zeros. $\mathbf{\Phi}_{t,s}$ is a lower Hessenberg matrix of order k (for a discussion of Hessenberg matrices, see Karanasos et al., 2022). It is also a banded matrix with total bandwidth $p+1$ (the number of its non-zero diagonals, i.e., the diagonals whose elements are not all identically zero), upper bandwidth 1 (the number of its non-zero super-diagonals), and lower bandwidth $p-1$ (the number of its non-zero sub-diagonals). In particular, the elements of $\mathbf{\Phi}_{t,s}$ are: (-1) occupying the entries of the superdiagonal, the values of the first autoregressive coefficient $\phi_1(\cdot)$ (from time $s+1$ to time t), occupying the entries of the main diagonal, the values of the $(1+r)$-th autoregressive coefficient $\phi_{1+r}(\cdot)$ for $r = 1, 2, \ldots, p-1$ (from time $s+1+r$ to time t), occupying the entries of the r-th sub-diagonal, and zero entries elsewhere. It is clear that for $p \geq k$, $\mathbf{\Phi}_{t,s}$ is a full lower Hessenberg matrix.

For every pair $(\tau, s) \in \mathbb{Z}^2$ with $s < \tau$, the principal determinant associated with eq. (41) is given by:

$$\xi(\tau, s) = \det(\mathbf{\Phi}_{\tau,s}). \quad (42)$$

Formally $\xi(\tau, s)$ is a lower Hessenbergian (determinant of a lower Hessenberg matrix; for details on Hessenbergians see, Karanasos et al. 2022). We further extend the definition of $\xi(\tau, s)$ so as to be defined over $\mathbb{Z}^2$ by assigning the initial conditions:

$$\xi(\tau, s) = \begin{cases} 1 & \text{if} \quad \tau = s \\ 0 & \text{if} \quad \tau < s. \end{cases} \quad (43)$$

For the theorem below the following notation is employed:

$$\xi^{(m)}(\tau, s) = \sum_{n=1}^{p+1-m} \phi_{m-1+n}(s+n)\xi(\tau, s+n). \quad (44)$$

The results on TV-DE equations in Karanasos et al. (2022), see theorem in their paper, are applied herein to obtain in the following theorem an equivalent explicit representation of the second moment process $\mu_{2\tau}$ in eq. (39) in terms of the prescribed variable: $\mu_{2s} = \mu_2$ (see eq. (30)).

Theorem 5 *An equivalent explicit representation of $\mu_{2\tau}$ in eq. (39) in terms of the prescribed variable μ_2 for any $s \in \mathbb{Z}$ and $\tau \in \mathbb{Z}_{s+1-p}$ such that $s < \tau$ is given by:*

$$\mu_{2\tau} = \mu_2 \sum_{m=1}^{p} \xi^{(m)}(\tau, s) + \sum_{l=s+1}^{\tau} \xi(\tau, l)c(l), \quad (45)$$

where $c(l)$, $\xi(\tau, l)$ and $\xi^{(m)}(\tau, s)$ are given in eqs. (38), (42)-(43) and (44), respectively.

Next we will consider case 1, where $\delta = 1$ and $\gamma(\tau) \neq 0$ for all τ.

Theorem 6 *The time varying second moment of σ_τ^2, that is $\mu_{2\tau}$, when $\delta = 1$ and $\gamma(\tau) \neq 0$ for all τ, obeys a time varying difference (TV-DE) equation of order p (with initial values μ_2) given by eq. (39) where $c(\tau)$ and $\phi_l(\tau)$ are given by*

$$\begin{aligned} c(\tau) &= \mu_{1\tau}^2 + \sum_{l=1}^{p} \mu_2\{\lambda_{41}(\tau, \tau-k) + \widetilde{\kappa}[\lambda_{42}(\tau, \tau-k) + \lambda_{43}(\tau, \tau-k)] + \kappa\lambda_{44}(\tau, \tau-k)\} \\ \phi_l(\tau) &= v_{41}(\tau, \tau-l) + \widetilde{\kappa}(\tau-l)[v_{42}(\tau, \tau-l) + v_{43}(\tau, \tau-l)] + \kappa(\tau-l)v_{44}(\tau, \tau-l). \end{aligned}$$

An equivalent explicit representation of $\mu_{2\tau}$ in eq. (39) in terms of the prescribed variable μ_2 for any $s \in \mathbb{Z}$ and $\tau \in \mathbb{Z}_{s+1-p}$ such that $s < \tau$ is given by eq. (45).

We skip the proof of theorem 6, which follows the same steps as the proofs of proposition 5 and theorem 5.

APPENDIX 4. Second-Order Persistence

In the following, we suggest a time varying second-order (or variance) persistence measure that is able to take into account the presence of breaks and to distinguish between the effects of a *mean shock* and a *volatility shock* on the level and conditional variance respectively. Fiorentini and Sentana (1998) argue that any reasonable measure of shock persistence should be based on the IRFs. For a univariate process x_t with *i.i.d* errors, e_t, they define the persistence of a shock e_t on x_t as $P(x_t\,|e_t) = \mathbb{V}ar(x_t)/\mathbb{V}ar(e_t)$. Clearly $P(x_t\,|e_t)$ will take its minimum value of one if x_t is white noise and it will not exist (will be infinite) for an $I(1)$, process.

A.4.1 Orthogonal Shocks

Recall that in general the shocks ε_t and v_t will be correlated with covariance matrix $\mathbf{\Sigma}_\tau$ (see eq. (4) in the main body of the paper). Next, we define the vector $\widetilde{\varepsilon}_\tau$ with two uncorrelated white noise shocks $\tilde{\varepsilon}_t$ and $\tilde{v}_t$ with variances equal to one. The relation between the original shocks and the orthogonal shocks is given by

$$\varepsilon_\tau = \widetilde{\mathbf{\Sigma}}_\tau \widetilde{\varepsilon}_\tau,$$

where

$$\widetilde{\mathbf{\Sigma}}_\tau = \begin{pmatrix} \sqrt{\sigma_{\varepsilon\tau}} & 0 \\ \rho_{\varepsilon v,\tau}\sqrt{\sigma_{v\tau}} & \sqrt{\sigma_{v\tau}}\sqrt{1-\rho_{\varepsilon v,\tau}^2} \end{pmatrix}. \tag{46}$$

It is straightforward to show that $\mathbf{\Sigma}_\tau{=}\widetilde{\mathbf{\Sigma}}_\tau\widetilde{\mathbf{\Sigma}}_\tau'$, which, by applying the vec operator, implies that

$$\mathbf{s}_\tau{=}\widetilde{\mathbf{\Sigma}}_\tau^{\otimes 2} vec(\mathbf{I}).$$

In view of eq. (46) the lower triangular matrix $\widetilde{\mathbf{\Sigma}}_\tau^{\otimes 2}$ is given by

$$\widetilde{\mathbf{\Sigma}}_\tau^{\otimes 2} = \begin{pmatrix} \sigma_{\varepsilon\tau} & 0 & 0 & 0 \\ \sigma_{\varepsilon\upsilon,\tau} & \sqrt{\sigma_{\varepsilon\tau}\sigma_{\upsilon\tau}(1-\rho^2_{\varepsilon\upsilon,\tau})} & 0 & 0 \\ \sigma_{\varepsilon\upsilon,\tau} & 0 & \sqrt{\sigma_{\varepsilon\tau}\sigma_{\upsilon\tau}(1-\rho^2_{\varepsilon\upsilon,\tau})} & 0 \\ \sigma_{\upsilon\tau}\rho^2_{\varepsilon\upsilon,\tau} & \rho_{\varepsilon\upsilon,\tau}\sigma_{\upsilon\tau}\sqrt{1-\rho^2_{\varepsilon\upsilon,\tau}} & \rho_{\varepsilon\upsilon,\tau}\sigma_{\upsilon\tau}\sqrt{1-\rho^2_{\varepsilon\upsilon,\tau}} & \sigma_{\upsilon\tau}(1-\rho^2_{\varepsilon\upsilon,\tau}) \end{pmatrix}. \tag{47}$$

Next we will present $\widetilde{\mathbf{\Sigma}}_\tau$ and $\widetilde{\mathbf{\Sigma}}_\tau^{\otimes 2}$ for the two cases 1 and 2.

Case 3 *When $\delta = 2$ and $\gamma = 0$, then $\widetilde{\Sigma}_\tau$ is a diagonal matrix given by:*

$$\widetilde{\mathbf{\Sigma}}_\tau = \begin{pmatrix} \sqrt{\mu_{1\tau}} & 0 \\ 0 & \sqrt{\mu_{2\tau}\kappa} \end{pmatrix}$$

(coincides with $\sqrt{\mathbf{\Sigma}_\tau}$. We recall that $\mu_{1\tau}$ has been derived in Section 3.2 in the main body of the paper and $\mu_{2\tau}$is given in eq. (45). Accordingly, $\widetilde{\mathbf{\Sigma}}_\tau^{\otimes 2}$ is given by

$$\widetilde{\mathbf{\Sigma}}_\tau^{\otimes 2} = diag(\mu_1 \ \sqrt{\mu_1\mu_2\kappa} \ \sqrt{\mu_1\mu_2\kappa} \ \mu_2\kappa). \tag{48}$$

Case 4 *When $\delta = 1$ and $\gamma \neq 0$, $\widetilde{\mathbf{\Sigma}}_\tau$ becomes*

$$\widetilde{\mathbf{\Sigma}}_\tau = \sqrt{\mu_{2\tau}} \begin{pmatrix} 1 & 0 \\ \widetilde{\kappa}(\tau) & \sqrt{\kappa(\tau) - \widetilde{\kappa}(\tau)^2} \end{pmatrix},$$

and, correspondingly:

$$\widetilde{\mathbf{\Sigma}}_\tau^{\otimes 2} = \mu_{2\tau} \begin{pmatrix} 1 & 0 & 0 & 0 \\ \widetilde{\kappa}(\tau) & \sqrt{\kappa(\tau) - \widetilde{\kappa}(\tau)^2} & 0 & 0 \\ \widetilde{\kappa}(\tau) & 0 & \sqrt{\kappa(\tau) - \widetilde{\kappa}(\tau)^2} & 0 \\ \widetilde{\kappa}(\tau)^2 & \widetilde{\kappa}(\tau)\sqrt{\kappa(\tau) - \widetilde{\kappa}(\tau)^2} & \widetilde{\kappa}(\tau)\sqrt{\kappa(\tau) - \widetilde{\kappa}(\tau)^2} & \kappa(\tau) - \widetilde{\kappa}(\tau)^2 \end{pmatrix}. \tag{49}$$

A4.2 Time Invariant Case

Next we will examine the case with constant coefficients. We reiterate that when the coefficients are constant $\boldsymbol{\gamma}$ is presented in theorem 2.

Theorem 7 *The second-order measure of persistence, that is $\mathbb{V}ar(y_\tau)$, for $\tau \leq t - k_2$, is decomposed in the persistence of the two orthogonal shocks, $\widetilde{\varepsilon}_t$ and $\widetilde{\upsilon}_t$, as follows*

$$\mathbb{V}ar(y_\tau) = P^{(Var)}(y_\tau \,|\widetilde{\varepsilon}) + P^{(Var)}(y_\tau \,|\widetilde{\upsilon}), \tag{50}$$

where

$$\begin{aligned} P^{(Var)}(y_\tau \,|\widetilde{\varepsilon}) &= g_{11}\sigma_\varepsilon + (g_{12} + g_{13})\sigma_{\varepsilon\upsilon} + g_{14}\sigma_\upsilon\rho^2_{\varepsilon\upsilon}, \\ P^{(Var)}(y_\tau \,|\widetilde{\upsilon}) &= g_{14}\sigma_\upsilon(1 - \rho^2_{\varepsilon\upsilon}). \end{aligned}$$

The superscript is used to distinguish the second-order measures of persistence from the first-order measures, see Section 3.3 in the main body of the paper. Hereafter for notational ease we will drop the superscript. Notice that when the two shocks are uncorrelated, $P(y_\tau \,|\widetilde{\varepsilon}\,)$ and $P(y_\tau \,|\widetilde{v}\,)$ in eq. (50) reduce to

$$P(y_\tau \,|\widetilde{\varepsilon}\,) = g_{11}\sigma_\varepsilon,\; P(y_\tau \,|\widetilde{v}\,) = g_{14}\sigma_v.$$

Proof. Replacing $\mathbf{s}$ in eq. (26) by $\widetilde{\mathbf{\Sigma}}^{\otimes 2} vec(\mathbf{I})$, on account of eq. (47) and using straightforward matrix algebra we obtain the first element of $\boldsymbol{\gamma}$, that is $\mathbb{V}ar(y_\tau)$, for $\tau \leq t - k_2$, as follows:

$$\mathbb{V}ar(y_\tau) = \underbrace{g_{11}\sigma_\varepsilon + (g_{12} + g_{13})\sigma_{\varepsilon v} + g_{14}\sigma_v\rho^2_{\varepsilon v}}_{P(y_\tau |\widetilde{\varepsilon}\,)} + \underbrace{g_{14}\sigma_v(1 - \rho^2_{\varepsilon v})}_{P(y_\tau |\widetilde{v}\,)},$$

and the proof is completed. ■

Next we will examine the two special cases 1 and 2.

Proposition 6 *When $\delta = 2$ and $\gamma = 0$, then*

$$P(y_\tau \,|\widetilde{\varepsilon}\,) = g_{11}\mu_1,\; P(y_\tau \,|\widetilde{v}\,) = g_{14}\mu_2\kappa.$$

When $\delta = 1$ and $\gamma \neq 0$, then

$$P(y_\tau \,|\widetilde{\varepsilon}\,) = \mu_2\{g_{11} + \widetilde{\kappa}[(g_{12} + g_{13}) + g_{14}\widetilde{\kappa}]\},\; P(y_\tau \,|\widetilde{v}\,) = g_{14}\mu_2(\kappa - \widetilde{\kappa}^2).$$

The proof is easy and therefore is omitted.

A4.3 Time Varying Case

We reiterate that when the coefficients are variable $\boldsymbol{\gamma}_\tau$ is derived in theorem 4.

Theorem 8 *The second-order measure of persistence, that is $\mathbb{V}ar(y_\tau)$, for $\tau > t - k_1$, is decomposed in the persistence of the two orthogonal shocks, $\widetilde{\varepsilon}_t$ and $\widetilde{v}_t$, as follows*

$$\mathbb{V}ar(y_\tau) = P(y_\tau \,|\widetilde{\varepsilon}\,) + P(y_\tau \,|\widetilde{v}\,),$$

where

$$\begin{aligned} P(y_\tau \,|\widetilde{\varepsilon}\,) &= \sum_{l=0}^{k_2+r-1} [v_{11}(\tau, \tau-l)\sigma_{\varepsilon,\tau-l} + (v_{12}(\tau, \tau-l) + v_{13}(\tau, \tau-l))\sigma_{\varepsilon v,\tau-l} + v_{14}(\tau, \tau-l)\sigma_{v,\tau-l}\rho^2_{\varepsilon v,\tau-l}] \\ &\quad +[\lambda_{11}(\tau, \tau-k)\sigma_\varepsilon + (\lambda_{12}(\tau, \tau-k) + \lambda_{13}(\tau, \tau-k))\sigma_{\varepsilon v} + \lambda_{14}(\tau, \tau-k)\sigma_v\rho^2_{\varepsilon v}], \qquad (51) \\ P(y_\tau \,|\widetilde{v}\,) &= \sum_{l=0}^{k_2+r-1} v_{14}(\tau, \tau-l)\sigma_{v,\tau-l}(1 - \rho^2_{\varepsilon v,\tau-l}) + \lambda_{14}(\tau, \tau-k)\sigma_v(1 - \rho^2_{\varepsilon v}). \qquad (52) \end{aligned}$$

Proof. Replacing in eq. (33) $\mathbf{s}_{\tau-l}$ by $\widetilde{\mathbf{\Sigma}}^{\otimes 2}_{\tau-l} vec(\mathbf{I})$ we obtain

$$\boldsymbol{\gamma}_\tau = \sum_{l=0}^{k_2+r-1} \mathbf{V}(\tau, \tau-l)\widetilde{\mathbf{\Sigma}}^{\otimes 2}_{\tau-l} vec(\mathbf{I}) + \mathbf{\Lambda}(\tau, \tau-k)\widetilde{\mathbf{\Sigma}}^{\otimes 2} vec(\mathbf{I}). \qquad (53)$$

On account of eqs. (31), (32), and (47), using straightforward matrix algebra the first element of $\boldsymbol{\gamma}_\tau$, that is $\mathbb{V}ar(y_\tau)$, for $\tau > t - k_1$, is given by

$$\mathbb{V}ar(y_\tau) = P(y_\tau\,|\widetilde{\varepsilon}\,) + P(y_\tau\,|\widetilde{v}\,),$$

where $P(y_\tau\,|\widetilde{\varepsilon}\,)$ and $P(y_\tau\,|\widetilde{v}\,)$ are defined in eqs. (51) and (52), respectively, and the proof is completed. ■

In the following proposition we will present the second-order persistence of the two orthogonal shocks, $\widetilde{\varepsilon}_t$ and $\widetilde{v}_t$, that is $P(y_\tau\,|\widetilde{\varepsilon}\,)$ and $P(y_\tau\,|\widetilde{v}\,)$, respectively, for the two cases 3 and 4.

Proposition 7 *Let $\delta = 2$ and $\gamma = 0$. Then*

$$\begin{aligned} P(y_\tau\,|\varepsilon\,) &= \sum_{l=0}^{k_2+r-1} v_{11}(\tau,\tau-l)\mu_{1,\tau-l} + \lambda_{11}(\tau,\tau-k)\mu_1, \\ P(y_\tau\,|v\,) &= k\left\{\sum_{l=0}^{k_2+r-1} v_{14}(\tau,\tau-l)\mu_{2,\tau-l} + \lambda_{14}(\tau,\tau-k)\mu_2\right\}. \end{aligned}$$

When $\delta = 1$ and $\gamma \neq 0$, $P(y_\tau\,|\widetilde{\varepsilon}\,)$ and $P(y_\tau\,|\widetilde{v}\,)$ are given by

$$\begin{aligned} P(y_\tau\,|\widetilde{\varepsilon}\,) &= \sum_{l=0}^{k_2+r-1} \mu_{2,\tau-l}\{v_{11}(\tau,\tau-l) + \widetilde{\kappa}(\tau-l)[(v_{12}(\tau,\tau-l) + v_{13}(\tau,\tau-l) + v_{14}(\tau,\tau-l)\widetilde{\kappa}(\tau-l)]\} \\ &\quad + \mu_2\{\lambda_{11}(\tau,\tau-k) + \widetilde{\kappa}[\lambda_{12}(\tau,\tau-k) + \lambda_{13}(\tau,\tau-k) + \lambda_{14}(\tau,\tau-k)\widetilde{\kappa}]\}, \\ P(y_\tau\,|\widetilde{v}\,) &= \sum_{l=0}^{k_2+r-1} v_{14}(\tau,\tau-l)[\kappa(\tau-l) - \widetilde{\kappa}(\tau-l)^2] + \lambda_{14}(\tau,\tau-k)(\kappa - \widetilde{\kappa}^2). \end{aligned}$$

In the absence of asymmetries, since $\widetilde{\kappa}(\tau) = 0$ for all τ, the above expressions reduce to

$$\begin{aligned} P(y_\tau\,|\widetilde{\varepsilon}\,) &= \sum_{l=0}^{k_2+r-1} \mu_{2,\tau-l}v_{11}(\tau,\tau-l) + \mu_2\lambda_{11}(\tau,\tau-k), \\ P(y_\tau\,|\widetilde{v}\,) &= \sum_{l=0}^{k_2+r-1} v_{14}(\tau,\tau-l)\kappa(\tau-l) + \lambda_{14}(\tau,\tau-k)\kappa. \end{aligned}$$

We skip proof of the above proposition (where we make use of eqs. (48) and (49)), which follows the same steps as the proof of theorem 8.

References

[1] AMISANO G., and R. GIACOMINI (2007): “Comparing Density Forecasts via Weighted Likelihood Ratio Tests”, *Journal of Business and Economic Statistics*, 25, 177–190.

[2] ANGELONI, I., COENEN, G. and F. SMETS (2003). "Persistence, the transmission mechanism and robust monetary Policy", *Scottish Journal of Political Economy*, 50, 527-549.

[3] APERGIS N. (2004): “Inflation, Output Growth, Volatility and Causality: Evidence from Panel Data and the G7 Countries”, *Economic Letters,* 83, 185-191.

[4] ATKESON, A., and L. OHANIAN (2001): “Are Phillips Curves Useful for Forecasting Inflation?”, *Federal Reserve Bank of Minneapolis Quarterly Review,* 25, 2-11.

[5] BAI, J., and P. PERRON (2003): “Computation and Analysis of Multiple Structural Change Models”, *Journal of Applied Econometrics*, 18, 1-22.

[6] BAILLIE, R., C. CHUNG, and M. TIESLAU (1996): “Analyzing Inflation by the Fractionally Integrated ARFIMA-GARCH Model”, *Journal of Applied Econometrics*, 11, 23-40.

[7] BAKER S., N. Bloom, S. J. Davis (2016): “Measuring Economic Policy Uncertainty”, *Quarterly Journal of Economics*, 131, 1593–1636.

[8] BALL, L., and S. MAZUMDER (2011): “Inflation Dynamics and the Great Recession”, *Brookings Papers on Economic Activity*, 42, 337-405.

[9] BARNETT W., F. JAWADI, and Z. FTITI (2020): “Causal Relationships Between Inflation and Inflation Uncertainty”, *Studies in Nonlinear Dynamics and Econometrics*, 24, 5-26.

[10] BATINI, N. (2006): “Euro Area Inflation Persistence”, *Empirical Economics,* 31, 977-1002.

[11] BENATI, L. (2008). "Investigating inflation persistence across monetary regimes". *The Quarterly Journal of Economics*, 123, 1005-1060.

[12] BLOOM, N. (2009): “The Impact of Uncertainty Shocks”, *Econometrica*, 77, 623-685.

[13] BOBEICA, E., and M. JAROCINSKI (2019): “Missing Disinflation and Missing Inflation: a VAR Perspective”, *International Journal of Central Banking,* 15, 199-232.

[14] BORN B., and J. PFEIFER (2014): “Policy Risk and the Business Cycle”, *Journal of Monetary Economics*, 68, 68-85.

[15] CANEPA, A., M. KARANASOS, and A. G. PARASKEVOPOULOS (2019): “Second Order Time Dependent Inflation Persistence in the United States: a GARCH-in-Mean Model with Time Varying Coefficients”, EST WP 11/19 University of Turin.

[16] BRAINARD, W., PERRY, G., (2000). Making Policy in a Changing World. In Perry, G., Tobin, J., eds., *Economic Events, Ideas, and Policies: The 1960s and After*. Washington, DC: Brookings Institution.

[17] CAPORALE, G. M., and A. KONTONIKAS (2009): “The Euro and Inflation Uncertainty in the European Monetary Union”, *Journal of International Money and Finance*, 28, 954-97.

[18] CAPORALE, G. M., L. ONORANTE, and P. PAESANI (2010): "Inflation and Inflation Uncertainty in the Euro Area", *ECB Working Papers Series* No 1229.

[19] CHANG K. L. (2012): "The Impacts of Regime-Switching Structures and Fat-Tailed Characteristics on the Relationship Between Inflation and Inflation Uncertainty", *Journal of Macroeconomics*, 34, 523-536.

[20] CLARIDA R., J. GALI, and M. GERTLER (2000): "Monetary Policy Rule and Macroeconomic Stability: Evidence and Some Theory", *Quarterly Journal of Economics,* 115, 147–180.

[21] COGLEY, T., and T. J. SARGENT (2001): "Evolving Post-World War II U.S. Inflation Dynamics," In Getler, M., and Rogoff, K. (Eds.), *NBER Macroeconomics Annual*, 331-373. MIT Press, Cambridge.

[22] COGLEY, T., and T. J. SARGENT (2005): "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US", *Review of Economic Dynamics*, 8, 262-302.

[23] CONRAD, C., and M. KARANASOS (2015a): "On the Transmission of Memory in GARCH-in-mean Models", *Journal of Time Series Analysis*, 36, 706-720.

[24] CONRAD, C., and M. KARANASOS (2015b): "Modeling the Link Between US Inflation and Output: the Importance of the Uncertainty Channel", *Scottish Journal of Political Economy*, 62, 431-453.

[25] CONRAD, C., M. KARANASOS, and N. ZENG (2010): "The Link Between Macroeconomic Performance and Variability in the UK", *Economics Letters,* 106, 154–157.

[26] CONRAD, C., and E. MAMMEN (2016): "Asymptotics for Parametric GARCH-in-mean Models", *Journal of Econometrics*, 194, 319-329.

[27] CORRADI, V., and N. SWANSON (2006): "Predictive Density Evaluation". In G. Elliott, C. Granger, and A. Timmermann (Eds.), *Handbook of Economic Forecasting*. North Holland.

[28] CUKIERMAN, A., and A. H. MELTZER (1986): "A Theory of Ambiguity, Credibility, and Inflation Under Discretion and Asymmetric Information", *Econometrica*, 54, 1099–1128.

[29] DAAL, E., A. NAKA, and B. SANCHEZ (2005): "Re-examining Inflation and Inflation Uncertainty in Developed and Emerging Countries", *Economics Letters* 89, 180-186.

[30] Del NEGRO. M., M. LENZA, G. PRIMICERI, and A. TAMABALOTTI (2020): "What's Up With the Phillips Curve", *Brookings Papers on Economic Activity*, Spring, 301-357.

[31] DIEBOLD, F., and R. MARIANO (1995): "Comparing Predictive Accuracy", *Journal of Business and Economics Statistics*, 13, 253-63.

[32] DING, Z., C. W. J. GRANER, and R. F. ENGLE (1993): "A Long Memory Property of Stock Market Returns and a New Model", *Journal of Empirical Finance*, 1, 83-106.

[33] ENGLE, R. F., D. M. LILIEN, and R. P. ROBINS (1987): "Estimating Time Varying Risk Premia in the Term Structure: the ARCH-M Model", *Econometrica*, 55, 391-407.

[34] ESER, F., P. KARADI, P. LANE, L. MORETTI, and C. OSBAT (2020): "The Phillips Curve at the ECB", *ECB Working Paper Series, No 2400.*

[35] EVANS, M. (1991): "Discovering the Link Between Inflation Rates and Inflation Uncertainty", *Journal of Money, Credit, and Banking*, 23, 169-84.

[36] EVANS, M., and P. WACHTEL (1993): "Inflation Regimes and Sources of Inflation Uncertainty", *Journal of Money, Credit and Banking,* 25, 475-511.

[37] FIORENTINI, G., and E. SENTANA (1998): "Conditional Means of Time Series Processes and Time Series Processes for Conditional Means", *International Economic Review*, 39, 1101-1018.

[38] FISHER, B., M. LENZA, H. PILL, and L. REICHLIN (2009): "Money and Monetary Policy in the Euro Area 1999-2006", *Journal of International Money and Finance*, 1138-1164.

[39] FOUNTAS, S., and M. KARANASOS (2007): "Inflation, Output Growth, and Nominal and Real Uncertainty: Empirical Evidence for the G7", *Journal of International Money and Finance*, 26, 229-250.

[40] FOUNTAS, S., M. KARANASOS, and J. KIM (2006): "Inflation Uncertainty, Output Growth Uncertainty, and Macroeconomic Performance", *Oxford Bulletin of Economics and Statistics*, 68, 319-343.

[41] FRIEDMAN, M. (1977): "Nobel Lecture: Inflation and Unemployment", *Journal of Political Economy*, 85, 451–472.

[42] GERLACH, S., and P. TILLMANN (2012): "Inflation Targeting and Inflation Persistence in Asia–Pacific", *Journal of Asian Economics*, 23, 360-73.

[43] GEWEKE J., and G. AMISANO (2011): "Optimal Prediction Pools", *Journal of Econometrics,* 164, 130-141.

[44] GIACOMINI A., and H. WHITE (2006): "Tests for Conditional Predictive Ability", *Econometrica*, 74, 1545-1578.

[45] GIORDANI, P. and P. SODERLIND (2003): "Inflation Forecast Uncertainty", *European Economic Review*, 47, 1037-1059.

[46] GRIER K.B., O.T. HENRY, N. OLEKALNS, and K. SHIELDS (2004): "The Asymmetric Effects of Uncertainty on Inflation and Output Growth", *Journal of Applied Econometrics,* 19, 551–565.

[47] HASSET K. A., and G. E. METCALF (1999): "Investment with Uncertain Tax Policy: Does Random Tax Policy Discourage Investment?", *Economic Journal*, 109, 72-393.

[48] HE, C. AND T. TERÄSVIRTA (1999): "Fourth Moment Structure of the GARCH(p, q) Process," *Econometric Theory*, 15, 824-846.

[49] HIGGS R. (1997): "Regime Uncertainty: Why the Great Depression Lasted So Long and Why Prosperity Resumed after the War", *Independent Review*, 1, 561-590.

[50] HOLLAND S. (1995): "Inflation and Uncertainty: Tests for Temporal Ordering", *Journal of Money Credit and Banking* 27, 827-837.

[51] HUBRICH, K., and F. SKUDELNY (2016): "Forecast Combination for Euro Area Inflation: a Cure in Times of Crisis?", *ECB Working Papers No 1972*.

[52] Hwang Y. (2001): "Relationship Between Inflation Rate and Inflation Uncertainty", *Economic Letters,* 73, 179–186.

[53] KARANASOS, M., and J. KIM (2006): "A Re-examination of the Asymmetric Power ARCH Model", *Journal of Empirical Finance*, 13, 113-128.

[54] KARANASOS, M., and S. SCHURER (2008): "Is the Relationship Between Inflation and its Uncertainty Linear?", *German Economic Review*, 9, 265-286.

[55] KARANASOS, M., and N. ZENG (2013): "Conditional Heteroskedasticity inMacroeconomics Data: UK Inflation, Output Growth and Their Uncertainties". In Hashimzade, N., and M.A. Thornton (Eds.), *Handbook of Research Methods and Applications in Empirical Macroeconomics*. Edward Elgar Publishing, Inc.

[56] KARANASOS, M. (1999): "The Second Moment and the Autocovariance Function of the Squared Errors of the GARCH Model," *Journal of Econometrics,* 90, 63-76.

[57] KARANASOS, M., PARASKEVOPOULOS, A., MAGDALINOS, A. and A. CANEPA (2022): "A Unified Theory for ARMA Models with Varying Coefficients: One Solution Fits All", unpublished paper.

[58] KIM C., C. NELSON, J. PIGER (2004). "The Less-Volatile U.S. Economy: A Bayesian Investigation of Timing, Breadth, and Potential Explanations", Journal of Business and Economic Statistics, 22, 80-93.

[59] KONTONIKAS, A. (2004): "Inflation and Inflation Uncertainty in the United Kingdom, Evidence from GARCH Modeling", *Economic Modelling,* 21, 525–543.

[60] KOOP, G., and S. M. POTTER (2007): "Estimation and Forecasting in models with Multiple Breaks", *The Review of Economic Studies*, 74, 763-789.

[61] LEVIN, A. T., and J. PIGER (2004): "Is Inflation Persistence Intrinsic in Industrial Economies?", *Working Paper 2002-023B*, Federal Reserve Bank of St. Louis.

[62] LUUKKONEN, R., P. SAIKKONEN, and T. TERASVIRTA (1988): "Testing Linearity Against Smooth Transition Autoregressive Models", *Biometrika*, 75, 491-499.

[63] MAHEU J.M., and S. GORDON (2008): "Learning, Forecasting and Structural Breaks", *Journal of Applied Econometrics*, 23, 553-583.

[64] McLEAY, M., and S. TENREYRO (2019): "Optimal Inflation and the Identification of the Phillips Curve", *NBER Macroeconomics Annual 2019*, Vol. 34.

[65] O'REILLY, G., K. WHELAN (2005): "Has Euro-area Inflation Persistence Changed Over Time?", *Review of Economics and Statistics,* 87, 709–720.

[66] ORPHANIDES, A., and S. VAN NORDEN (2005): "The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time", *Journal of Money, Credit and Banking*, 37, 583–601.

[67] PASTOR L., and P. Veronesi (2012): "Uncertainty about Government Policy and Stock Prices", *Journal of Finance*, 67, 1219-1264.

[68] PERRON, P., Y. YAMAMOTO, and J. ZHOU (2020): "Testing Jointly for Structural Changes in the Error Variance and Coefficients of a Linear Regression Model", *Quantitative Economics*, 11, 1019-1057.

[69] PESARAN, M. H., and A. TIMMERMANN (2005): "Small Sample Properties of Forecasts from Autoregressive Models under Structural Breaks", *Journal of Econometrics*, 129, 183-217.

[70] PESARAN, M. H., D. PETTENUZZO, and A. TIMMERMANN (2006): "Forecasting Time Series Subject to Multiple Structural Breaks", *The Review of Economic Studies*, 73, 1057-1084.

[71] PETTENUZZO, D., and A. TIMMERMAN (2017): "Forecasting Macroeconomic Variables Under Model Instability", *Journal of Business and Economic Statistics*, 35, 183-201

[72] PIVETTA, F., and R. REIS (2007): "The Persistence of Inflation in the United States", *Journal of Economic Dynamics and Control*, 31, 1326-1358.

[73] RODRIK D. (1991): "Policy Uncertainty and Private Investment", *Journal of Development Economics*, 36, 229-242.

[74] ROSSI B., and T. SEKHPOSYAN (2010): "Have Economic Models' Forecasting Performance for US Output Growth and Inflation Changed Over Time, and When?", *International Journal of Forecasting*, 26, 808–835.

[75] SCOTTI C. (2016): "Surprise and Uncertainty Indexes: Real-Time Aggregation of Real-Activity Macro Surprises", *Journal of Monetary Economics*, 82, 1-9.

[76] SENSIER, M., and D. VAN DIJK (2004): "Testing for Volatility Changes in US Macroeconomic Time Series", *Review of Economics and Statistics*, 86, 833–839.

[77] STOCK, J. H. (2001): "Comment on: Evolving Post-World War II U.S. Inflation Dynamics", In: Gertler, M., and Rogoff, K. (Eds.), *NBER Macroeconomics Annual*, 379-387. MIT Press, Cambridge.

[78] STOCK, J. H., and M. W. WATSON (2007): "Why has U.S. inflation Become Harder to Forecast?", *Journal of Money, Credit and Banking*, 39, 3–33.

[79] STOCK, J.H., and M. W. WATSON (2009): "Phillips Curve Inflation Forecasts", Ch. 3 in Fuhrer, J., Y. Kodrzycki, Little, J., and Olivei, J. (Eds.), *Understanding Inflation and the Implications for Monetary Policy*, 101-186, Cambridge: MIT Press (with M. W. Watson).

[80] TAYLOR, J. B. (2000): "Low Inflation Pass-Through, and Pricing Power of Firms", *European Economic Review,* 44, 1389-1408.

[81] TERASVIRTA, T., D. VAN DIJK, and M. C. MEDEIROS (2005): "Linear Models, Smooth Transition Autoregressions, and Neural Networks for Forecasting Macroeconomic Time Series: A Re-Examination", *International Journal of Forecasting*, 21, 755-774.

[82] WILSON B.K. (2006): "The Links Between Inflation, Inflation Uncertainty and Output Growth: new time series evidence from Japan", *Journal of Macroeconomics* 28, 609-620.