Density forecast transformations

Matteo Mogliani[1] and Florens Odendahl[2]

[1]*Banque de France* *

[2]*Banco de España*†

PRELIMINARY DRAFT — DO NOT CIRCULATE WITHOUT EXPLICIT PERMISSION
February 23, 2024

Abstract

The popular choice of using a *direct* forecasting scheme implies that the individual predictions do not contain information on cross-horizon dependence. However, this dependence is needed if the forecaster has to construct, based on the *direct* forecasts, predictive objects that are functions of several horizons; such as obtaining annual-average from quarter-on-quarter growth rates. To address this issue we propose to use copulas to combine the individual h-step-ahead predictive distributions into a joint predictive distribution. Our method is particularly appealing for practitioners for whom changing the *direct* forecasting specification is too costly. In a Monte Carlo study, we demonstrate that our approach leads to a better approximation of the true predictive densities than an approach which ignores the potential dependence. We show the superior performance of our method in several empirical examples, where we construct (i) quarterly forecasts using month-on-month *direct* forecasts, (ii) annual-average forecasts using monthly year-on-year *direct* forecasts, and (iii) annual-average forecasts using quarter-on-quarter *direct* forecasts.

Keywords: joint predictive distribution, frequency transformation, path forecasts, cross-horizon dependence.
JEL codes: C53, C32, E37.

*Address: 31 Rue Croix des Petits Champs, 75001 Paris, France. Email: matteo.mogliani@banque-france.fr.

†Address: Calle de Alcalá 48, 28014 Madrid, Spain. Email: florens.odendahl@bde.es. We thank Mohammed Chahad and participants at the 12th ECB Conference on Forecasting Techniques and the Conference on Real-Time Data Analysis, Methods, and Applications for insightful discussions. The paper previously circulated under the title "Density forecast frequency transformation via Copulas". The views expressed herein are those of the authors and should not be attributed to the Banco de España, the Banque de France or the Eurosystem.

1 Introduction

Forecasting models are often specified to produce *direct* h-step-ahead forecasts, which implies that the predictions do not contain information on their cross-horizon dependence. As a consequence the individual h-step-ahead predictions cannot easily be transformed into predictions of objects that depend on several horizons. For instance, the literature on macroeconomic risk uses quantile models that, in brief, produce *direct* density forecasts of quarter-on-quarter (qoq) real GDP growth (Adrian et al., 2019; Ferrara et al., 2022); however, it is not straightforward to make additional use of the quarter-on-quarter density forecasts by constructing annual-average growth rates.

To address this issue, we propose using Gaussian copulas to combine the information of *direct* h-step-ahead predictive densities into a joint distribution. This allows the practitioner to construct predictive objects that are functions of several horizons, which we label target-frequency predictive densities, since the multivariate distribution of the direct h-step-ahead predictions reflects the serial dependence between the forecasts.[1]

The scenario of having a fixed forecasting specification is particularly common in institutions, such as central banks, where changing the forecasting process is costly and yet transformations of the existing forecasts to other target-frequencies are often required. Further, for some target-frequencies the available time series data can be too short for estimation. This is the case, for instance, if the target-frequency is in annual-averages of calendar years, a target-frequency often reported in institutional forecasts.[2] Another reason can be that the forecasts are derived from surveys which only report one frequency but not the required target-frequency. In general, our approach helps to broaden the usability of already available individual predictive densities that are based on a *direct* forecasting scheme.

As an alternative to our approach, the researcher could use a simple approach that assumes independence between the different marginal predictive densities, i.e., no correlation between the direct h-step-ahead predictions at different horizons. However, this implies to ignore the serial dependence that is typically present in macroeconomic variables. We show in several Monte Carlo studies that our approach delivers better approximations to the true underlying annual-average forecasts for different DGPs, even under misspecified forecasting models, when the true multivariate distribution is not Gaussian, and for fairly small training samples for the copula parameter estimation.

For the application of our approach, the researcher only needs to compute the correlation between the empirical PITs of the individual h-step-ahead predictive distributions for different horizons in a training sample. In particular, the forecaster needs to (i) compute the sequence of realized PITs for the marginal predictive densities at each forecast horizon h=1,..,H, from a pseudo out-of-sample exercise over a training sample and to (ii) combine the marginal distributions into a joint distribution via a multivariate Gaussian copula through the rank correlation estimated on the realized PITs.

We demonstrate the usefulness of our methodology in three empirical applications. The first empirical application is a large-scale forecasting exercise based on monthly data from

[1] Note that for models that produce *iterative* h-step-ahead predictive densities, for instance Vector Autoregressions, forecasts conditional on specific paths or the annual-average frequency transformation are in general possible since the iterative approach allows to draw conditional on previous horizons.

[2] See, for instance, the Macroeconomic projection of the ECB.

FRED-MD (McCracken and Ng, 2016), in which we first compute density forecasts for month-on-month values for a large number of randomly selected outcome variable and predictor combinations. We then use these predictive densities to compute quarter-on-quarter density forecasts through our proposed copula approach. Results show that the copula approach outperforms the approach that ignores cross-horizon dependence for the majority of outcome variable and predictor combinations.

The second empirical application aims to emulate a situation in which a forecaster has predictive densities for year-on-year inflation but to get a more complete picture of the inflation environment, the forecaster would like to extend their set of forecasts to include annual-average predictive densities. Importantly, the year-on-year and annual-average predictions need to be coherent, i.e., they should be based on the same predictors, model type, and the central tendency of the forecasts across the two frequencies should be very similar. Therefore, we provide *direct* predictive densities for year-on-year inflation and transform them, via the proposed copula method, into annual-average inflation based on the U.S. The copula approach provides significantly better forecasts of annual-average inflation than the benchmark approach, in particular at the tails of the distribution.

In the third, we use the predictive densities of quarter-on-quarter U.S. real GDP growth from Adrian et al. (2019), which are based on *direct* forecasts, and transform them into annual-average forecasts. In particular, our approach allows us to use the original predictive densities of Adrian et al. (2019) as the basis for our annual-average forecasts. We find that the annual-average forecasts based on the copula-approach leads to more accurate density forecasts than the benchmark approach.

Section 2 describes the methodological framework. Section 4 shows Monte Carlo results for the frequency transformation via copulas. Section 5 shows the three empirical exercises and Section 6 concludes.

2 Motivating example

Consider the following simple mean-zero autoregressive model:

$$Y_{t+1} = \rho Y_t + \varepsilon_{t+1}$$

with $|\rho| < 1$ and $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. It is well known that the optimal h-steap ahead prediction (under both iterated and direct forecasting approach) is given by

$$Y_{t+h|t} = \rho^h Y_t$$

It follows that the forecast error $e_{t+h|t} = \sum_{j=0}^{h-1} \rho^j \varepsilon_{t+h-j}$ has second moment:

$$\mathbb{V}(e_{t+h|t}) = \sigma_\varepsilon^2 \left(\frac{1-\rho^{2h}}{1-\rho^2} \right)$$

and auto-covariance and auto-correlation functions:

$$Cov(e_{t+h|t}, e_{t+h-k|t}) = \sigma_\varepsilon^2 \rho^k \left(\frac{1-\rho^{2(h-k)}}{1-\rho^2} \right)$$
$$Corr(e_{t+h|t}, e_{t+h-k|t}) = \rho^k \sqrt{\frac{1-\rho^{2(h-k)}}{1-\rho^{2h}}}$$

for $h > k > 0$. Note that *Corr* denotes the Pearson correlation. Thus, the process has conditional predictive distribution:

$$p(Y_{t+h|t}; \rho, \sigma_\varepsilon) \sim \mathcal{N}\left(\rho^h Y_t, \sigma_\varepsilon^2 \frac{1-\rho^{2h}}{1-\rho^2}\right)$$

Now consider a linear transformation of the forecast sequence $\{Y_{t+j|t}\}_{j=1}^h$, such as $Z_{\mathbf{t+h}|t} = c_1 Y_{t+1|t} + \cdots + c_h Y_{t+h|t}$. This transformation is often useful in macroeconomic applications when the original forecasts need to be converted into a different target periodic measure of the same phenomenon. For instance, if $Y_{t+h|t}$ is the h-step ahead forecast of a month-on-month growth rate, then for $h = 12$ and $c_1 = \cdots = c_h = 1$, the transformed forecast

$$Z_{\mathbf{t+h}|t} = \sum_{j=1}^{h} c_j Y_{t+j|t}$$

is (approximately) the 12-months ahead forecast of the year-on-year growth rate. In this case, we shall denote $Z_{\mathbf{t+h}|t}$ a *periodic-transformation* of the forecast sequence $\{Y_{t+j|t}\}_{j=1}^h$. Similarly, a linear transformation can map the forecasts generated at the original sampling frequency of the data into a forecast sequence sampled at a desired lower target frequency. For instance, for $h = 3$ and assuming that $Y_{t+1|t}$ is the forecast of a month-on-month growth rate for the first month of a given quarter, the transformed forecast

$$Z_{\mathbf{t+h}|t} = \frac{1}{3} Y_{t-1} + \frac{2}{3} Y_t + Y_{t+1|t} + \frac{2}{3} Y_{t+2|t} + \frac{1}{3} Y_{t+3|t}$$

is (approximately) the 3-months ahead forecast of the quarter-on-quarter growth rate (see Mariano and Murasawa, 2003).[3] In this case, we shall denote $Z_{\mathbf{t+h}|t}$ a *frequency-transformation* of the forecast sequence $\{Y_{t+j|t}\}_{j=1}^h$.

For the sake of simplicity, let's set $c_j = 1$, for $j = 1, \ldots, h$. The forecast error $e(Z)_{\mathbf{t+h}|t}$ is given by:

$$e(Z)_{\mathbf{t+h}|t} = \sum_{j=0}^{h-1} \frac{1-\rho^{j+1}}{1-\rho} \varepsilon_{t+h-j}$$

with second moment:

$$\mathbb{V}(e(Z)_{\mathbf{t+h}|t}) = \sigma_\varepsilon^2 \sum_{j=0}^{h-1} \left(\frac{1-\rho^{j+1}}{1-\rho} \right)^2.$$

[3] This formula can be generalised to other frequency transformations, such as from month-on-month or quarter-on-quarter growth rates to annual average growth rates, by changing the sequence of c_j.

Hence, the conditional predictive distribution of the "ideal" transformed forecast is:

$$p(Z_{\mathbf{t+h}|t};\rho,\sigma_\varepsilon) \sim \mathcal{N}\left(\sum_{j=0}^{h-1}\rho^{j+1}Y_t, \sigma_\varepsilon^2\sum_{j=0}^{h-1}\left(\frac{1-\rho^{j+1}}{1-\rho}\right)^2\right) \tag{1}$$

Note that this is equivalent to:

$$p(Z_{\mathbf{t+h}|t};\rho,\sigma_\varepsilon) \sim \mathcal{N}\left(\sum_{j=0}^{h-1}\rho^{j+1}Y_t, \sigma_\varepsilon^2\sum_{j=1}^{h}\left(\frac{1-\rho^{2j}}{1-\rho^2}\right) + 2\sigma_\varepsilon^2\sum_{j=2}^{h}\sum_{k=1}^{j-1}\rho^k\left(\frac{1-\rho^{2(j-k)}}{1-\rho^2}\right)\right),$$

which is the sum of random variables from the joint Multivariate-Normal forecast distribution

$$p(Y_{t+1|t},\ldots,Y_{t+h|t};\rho,\sigma_\varepsilon) \sim \mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$$

where

$$\boldsymbol{\mu} = \begin{bmatrix}\rho Y_t\\ \rho^2 Y_t\\ \rho^3 Y_t\\ \vdots\\ \rho^h Y_t\end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \sigma_\varepsilon^2\begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{h-1}\\ \rho & \left(\frac{1-\rho^4}{1-\rho^2}\right) & \rho\left(\frac{1-\rho^4}{1-\rho^2}\right) & \cdots & \rho^{h-2}\left(\frac{1-\rho^4}{1-\rho^2}\right)\\ \rho^2 & \rho\left(\frac{1-\rho^4}{1-\rho^2}\right) & \left(\frac{1-\rho^6}{1-\rho^2}\right) & \cdots & \rho^{h-3}\left(\frac{1-\rho^6}{1-\rho^2}\right)\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \rho^{h-1} & \rho^{h-2}\left(\frac{1-\rho^4}{1-\rho^2}\right) & \rho^{h-3}\left(\frac{1-\rho^6}{1-\rho^2}\right) & \cdots & \left(\frac{1-\rho^{2h}}{1-\rho^2}\right)\end{pmatrix}$$

The predictive distribution in (1) can be compared to the predictive distribution of the "inattentive" forecaster. This forecaster would typically ignore the cross-horizon dependence of the forecasts, *i.e.* the correlation structure of the forecast errors, and would draw predictions from the convoluted conditional forecast distribution:

$$p(\widetilde{Z}_{\mathbf{t+h}|t};\rho,\sigma_\varepsilon) \sim \mathcal{N}\left(\sum_{j=0}^{h-1}\rho^{j+1}Y_t, \sigma_\varepsilon^2\sum_{j=1}^{h}\left(\frac{1-\rho^{2j}}{1-\rho^2}\right)\right). \tag{2}$$

From (1) and (2), it is clear that the ideal and the inattentive forecast distributions diverge with $|\rho| \to 1$ and $h\uparrow$. Further, noting that the conditional data density is given by

$$p(Z^*_{\mathbf{t+h}};\rho,\sigma_\varepsilon) \sim \mathcal{N}\left(\sum_{j=0}^{h-1}\left(\rho^{j+1}Y_t + \frac{1-\rho^{j+1}}{1-\rho}\varepsilon_{t+h-j}\right), \frac{\sigma_\varepsilon^2}{1-\rho^2}\left(h + \sum_{j=1}^{h-1}2(h-j)\rho^j\right)\right),$$

it can be shown that the data and the ideal forecast distributions tend to converge with $h\uparrow$. These features are illustrated in the first panel of Figure 1, which shows that the Kullback-Leibler (KLIC) difference between the inattentive and the ideal forecasts, both computed with respect to the conditional true data density, is a monotonically increasing function of ρ and h.[4]

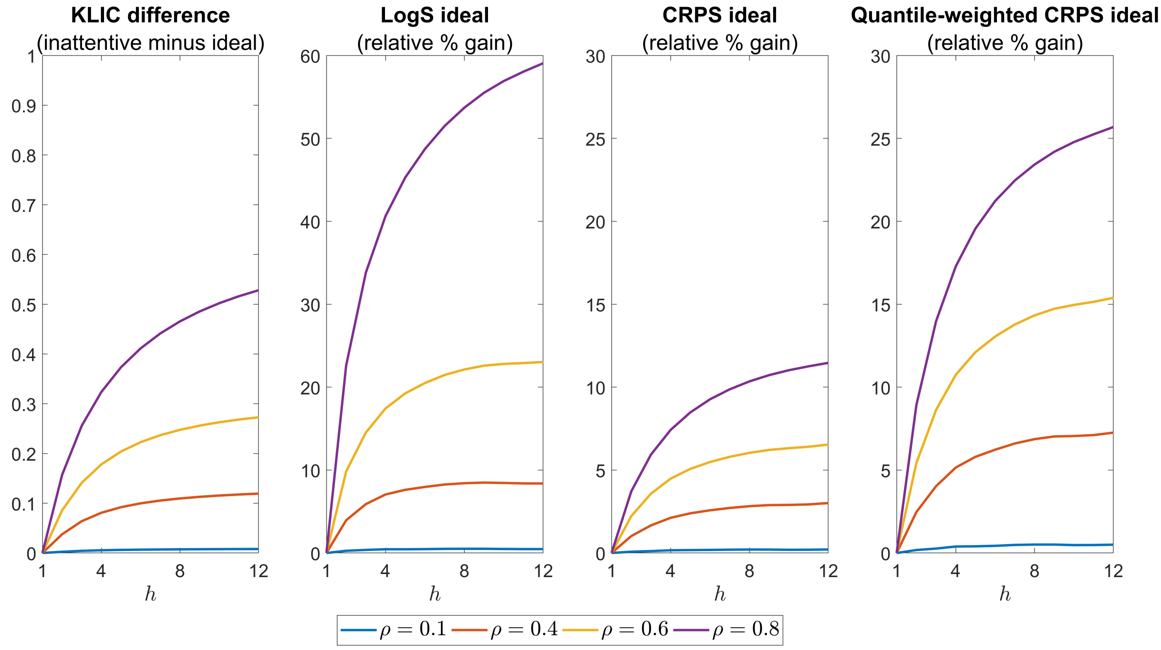
The remaining panels of Figure 1 show the average relative accuracy of the ideal density

[4] For Gaussian densities, the KLIC difference can be approximated by:

$$\text{KLIC}(\widetilde{Z},Z^*) - \text{KLIC}(Z,Z^*) \approx \log_e\left(\frac{\sigma_Z}{\sigma_{\widetilde{Z}}}\right) + \frac{\sigma^2_{\widetilde{Z}} - \sigma^2_Z}{2\sigma^2_{Z^*}}$$

where $\sigma^2_{\widetilde{Z}}$ and $\sigma^2_{Z^*}$ denote here, respectively, the variance of the inattentive and ideal density forecasts.

Figure 1: Scores for density forecasts: ideal *vs* inattentive approach



Note: KLIC difference denotes the difference between $\text{KLIC}(Z, \widetilde{Z})$ and $\text{KLIC}(Z, Z^*)$. LogS denotes the logarithmic score. CRPS denotes the continuous ranked probability score. Quantile-weigthed CRPS denotes the quantile weighted versions of the continuous ranked probability score, with emphasis on the tails. LogS, CRPS, and quantile-weigthed CRPS are expressed in relative % gain of the ideal forecaster with respect to the inattentive forecaster.

forecast compared to the inattentive forecast, evaluated through proper scoring rules, such as the logarithmic score (LogS), the continuous ranked probability score (CRPS), and the quantile-weigthed CRPS (qwCRPS), the latter with emphasis on the tails. All these metrics show robust gains for the ideal forecast, which increases monotonically with ρ and h. For instance, with $\rho = 0.6$ and $h = 12$ (the year-on-year growth rate transformation), the gain would stand about 20-25% according to the LogS and 7% according to the CRPS, while the qwCRPS points to a gain of about 15%. Not surprisingly, the latter suggests that important accuracy gain can be obtained when focusing on the tails of the predictive densities, rather than on their central part. This is due to the fact that the two densities differ solely on their variance, while the other moments are the same.[5]

3 Constructing multivariate densities with copulas

From the previous section, it is clear that constructing the correct predictive density of transformed forecasts requires drawing from the joint predictive distribution of the original marginal forecasts. However, this can be often impractical in empirical applications, such as those relying on direct multi-step forecasting or, more generally, when the marginal forecast distributions are not necessarily Gaussians.

To address this issue, in this paper we propose to resort on (Gaussian) copulas. For instance,

[5] It is nevertheless worth noting that these results depend on the sign of ρ. With $\rho < 0$, the monotonicity feature is in part lost, in particular for large (negative) autoregressive coefficients and small h. However, the ideal forecaster outperforms the inattentive forecaster even under this parameterization, in particular for large forecast horizons. These results are not presented, but are available upon request to the authors.

continuing on the example in Section 2, the predictive density $p(Z_{\mathbf{t+h}|t};\rho,\sigma_\varepsilon)$ can be constructed by drawing from the joint forecast distribution $p(Y_{t+1|t},\ldots,Y_{t+h|t};\rho,\sigma_\varepsilon)$. Defining $\Phi(Y_{t+h|t})$ the CDF of the Gaussian predictive distributions, we note that:

$$Corr\left(\Phi(Y_{t+h|t}),\Phi(Y_{t+h-k|t})\right)=\frac{6}{\pi}\arcsin\left(\frac{1}{2}\rho^k\sqrt{\frac{1-\rho^{2(h-k)}}{1-\rho^{2h}}}\right) \tag{3}$$

which is approximately the Spearman correlation coefficients of the PITs. The Gaussian copula is given by:

$$C_R=\Phi_R\left(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_h)\right)$$

where Φ^{-1} denotes the inverse CDF of a standard Normal and Φ_R the joint CDF of a standard multivariate Normal with covariance matrix R. Note that R is hence the correlation matrix, whose elements are defined in (3). Note also that the predictive distribution of the "inattentive" forecaster in (2) is also equivalent to the joint distribution of the forecasts constructed through a Gaussian copula, but with $R=\mathbf{I}_h$. Given the copula, it is easy to resample $Y_{t+1|t},\ldots,Y_{t+h|t}$ from their joint distribution:

$$\left(Y_{t+1|t},\ldots,Y_{t+h|t}\right)=\left(\Phi^{-1}(U_1),\ldots,\Phi^{-1}(U_h)\right)$$

and then compute the desired (periodic or frequency) transformed density forecast from the sampled joint forecasts.

We can now extend this approach to a more general forecasting environment. Assume the forecaster has a set of *direct* h-step-ahead predictive densities for T forecast origins, denoted by $\{\{g_{t,h}\}_{h=1}^H\}_{t=1}^T$ and with predictive cumulative distribution functions (cdf) $\{\{G_{t,h}\}_{h=1}^H\}_{t=1}^T$, for outcome variable Y_{t+h}; the subscript h denotes the forecast horizon and the subscript t denotes the forecast origin. Further assume that the set of predictive distributions, $\{\{g_{t,h}\}_{h=1}^H\}_{t=1}^T$, is taken as given, for instance, due to institutional restrictions on the forecasting model to be used.

To illustrate the application of our methodology, but without loss of generality, we will assume that the predictive density $g_{t,h}$ is a predictive density for quarter-on-quarter growth rates.[6] In period T, the forecaster is asked to provide predictive densities for the annual-average growth rates as well as for the conditional predictive density $\tilde{g}_{T,h}(y_{T+h}|y_{T+h-1},...,y_{T-1})$, henceforth called path-forecast, based on $\{\{g_{t,h}\}_{h=1}^H\}_{t=1}^T$.

We propose to do this by using copula functions, developed by Sklar (1959). A copula can be described as a function such that for any $Q(y_1,...,y_d)$, where Q is the multivariate distribution function of the random vector $(Y_1,...,Y_d)$, there is a copula function $C(\cdot|R)$, such that $Q(y_1,...,y_d)$ = $C(G_{Y_1}(y_1),...,G_{Y_d}(y_d)|R)$, where $G_{Y_1},...,G_{Y_d}$ are the marginal cdfs of $Y_1,...,Y_d$, respectively, and R denotes the parameter(s) that governs the dependence between $G_{Y_1}(y_1),...,G_{Y_d}(y_d)$. Inversely, a copula function C, combined with marginal cdfs $G_{Y_1},...,G_{Y_d}$, gives a multivariate distribution.

A popular copula family is the Gaussian copula, denoted by C_{Ga}, where the dependence between the d variables is governed by the correlation matrix R, with ones on the diagonal and the rank correlation of variable i and j as the respective off-diagonal element (i,j).

Let then $Q_T(y_{T+1},...,y_{T+h}|R)$ denote the joint predictive cdf of $Y_{T+1},...,Y_{T+h}$ for forecast origin

[6] Predictive distributions for month-on-month growth rates, monthly or quarterly (log-)levels or year-on-year growth rates can be handled analogously with our approach.

T, conditional on the correlation matrix R and constructed using C_{Ga}. Note that $Q_T(y_{T+1},...,y_{T+H}|R) = C_{\text{Ga}}(G_{T,1}(y_{T+1}),...,G_{T,H}(y_{T+H})|R)$. Let further $\text{PIT}_{t,h} = G_{t,h}(y_{t+h})$, where y_{t+h} is the realized value. The forecaster can obtain an estimate of $Q_T(y_{T+1},...,y_{T+H}|R)$ using the algorithm described below.

Algorithm 1 ***Joint Predictive Distribution***

1. *Compute the realized PITs,* $\{\{PIT_{t,h}\}_{h=1}^{H}\}_{t=1}^{T-H}$, *of the predictive CDFs* $\{\{G_{t,h}\}_{h=1}^{H}\}_{t=1}^{T-H}$.
2. *Compute the rank correlations of* $PIT_{t,h}$ *across the different* h *to get an estimate of* $\widehat{R}$.
3. *Use* $\widehat{R}$ *in combination with* C_{Ga} *to obtain the joint distribution* $\widehat{Q}_T(y_{T+1},..,y_{T+H}|\widehat{R})$.

The resulting multivariate distribution allows to sample the *direct* h-step-ahead predictions jointly, such that predictive objects that are functions of several horizons can be constructed.

To illustrate the use of Algorithm 1, consider the following example. The forecaster is asked in T, which is the last quarter of the year, to provide a predictive distribution of the annual-average growth for the next year. The forecaster, however, has only a set of *direct* quarter-on-quarter h-step-ahead growth rate predictions available, for $h = 1,...,4$. To transform the quarter-on-quarter growth rates into annual-average predictions, the forecaster can obtain the set $\{[Y_{T,1,s}, Y_{T,2,s}, Y_{T,3,s}, Y_{T,4,s}]\}_{s=1}^{S}$ of draws, for $s = 1,...,S$, from $\widehat{Q}_T^{-1}(y_{T+1},..,y_{T+4}|\widehat{R})$. This can easily be implemented in standard statistics software. First, compute the lower Cholesky decomposition of $\widehat{R}$, denoted by P. Then, for each $s = 1,...,S$:

(a) Draw a vector $X = [X_1,...,X_4]'$ of independent standard Normals.

(b) Compute the vector $U = [U_1, U_2, U_3, U_4] = [\Phi(Z_1), \Phi(Z_2), \Phi(Z_3), \Phi(Z_4)]$, where $\Phi(\cdot)$ is the CDF of a standard Normal distribution and $[Z_1, Z_2, Z_3, Z_4] = Z = PX$.

(c) Evaluate $G_{T,1}(U_1),...,G_{T,4}(U_4)$ to get the joint draw $[Y_{T,1,s}, Y_{T,2,s}, Y_{T,3,s}, Y_{T,4,s}]$.

The joint draws $\{[Y_{T,1,s}, Y_{T,2,s}, Y_{T,3,s}, Y_{T,4,s}]\}_{s=1}^{S}$ can be used to obtain the predictive distribution of the annual-average growth rate by computing the level of quarter two, three, and four for each $s = 1,...,S$.

The multivariate distribution also allows to sample $Y_{t,h,i}$ conditional on $Y_{t,h-j}$ for $j = 1,...,h-1$.

4 Monte Carlo study

We study the performance of our suggested copula approach via Monte Carlo simulations in a scenario where the forecaster has a model that produces *direct* h-step-ahead predictive densities for qoq growth rates and then needs to transform the predictive densities into annual-average growth rates and year-on-year (yoy) growth rates. The absolute and relative performance of the proposed approach with respect to a simple benchmark ignoring cross-horizon dependence of the forecasts (the "inattentive" forecaster described in Section 2) are evaluated via an out-of-sample forecasting exercise.

4.1 Monte Carlo design

*** Note that the results in the submitted version of the paper uses the value of $\theta_2 = 0$ from an earlier version of the paper. An update of the results with $\theta \neq 0$is work in progress ***

The underlying DGP of qoq growth rates, denoted by Y_t, takes the form of a VAR(1):

$$\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} + \begin{bmatrix} \theta_1 & \theta_2 \\ 0 & \gamma \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \tag{4}$$

where $\{\varepsilon_{j,t}\}_{t=1}^{T}$ are two uncorrelated sequences of independent and identically distributed (iid) shocks. We set $\varepsilon_{2,t} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \sigma^2_{\varepsilon_2}\right)$, but we consider three different specifications for the error term $\varepsilon_{1,t}$: (*i*) a Normal distribution, (*ii*) a Skew-Normal (SN) distribution, or (*iii*) a Skew-*t* (ST) distribution. For cases (*ii*) and (*iii*), we adopt a location-scale-shape parameterization (Azzalini and Capitanio, 2003). All the distributions are calibrated to have mean zero and standard deviation $\sigma_{\varepsilon_1} = 0.5$, as well as negative skewness for cases (*ii*) and (*iii*) (with shape parameter $\alpha = -1.5$). For the ST distribution, the degrees of freedom parameter is set to $\nu = 8$, which implies somewhat heavier tails than the Normal distribution. We consider these different specifications to allow for a varying degree of complexity in the DGP. In addition to the qoq growth rates, we also simulate the true annual-average and year-on-year growth rates generated by the DGP described in eq. (4).

We consider two types of forecasting models, both of which are misspecified AR(1) and produce *direct* h-step-ahead forecasts. The first forecasting model is used when $\varepsilon_{1,t}$ in the DGP is drawn from a Normal distribution:

$$Y_{t+h} = \tau_h + \beta_h Y_t + u_{t+h}, \tag{5}$$

with $u_{t+h} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2_{u,h})$.

The second forecasting model is a quantile regression specification used when $\varepsilon_{1,t}$ in the DGP is drawn from the Skew-Normal or the Skew-*t* distribution:

$$Y_{t+h}(q) = \tau_h(q) + \beta_h(q) Y_t + u_{t+h}(q), \tag{6}$$

where $q \in Q = [0.05, 0.25, 0.5, 0.75, 0.95]$ denotes the set of estimated quantiles, and $\tau_h(q)$ and $\gamma_h(q)$ denote respectively the quantile specific intercept and autoregressive parameter. To obtain a full predictive distribution, we smooth the five predicted quantiles using the Skew-*t* of Azzalini and Capitanio (2003). Note that the latter introduces a second potential source of model misspecification, in addition to that implied by the specification in (6).

The parameters $\{\tau_h, \beta_h\, \sigma^2_{u,h}\}$ and $\{\tau_h(q), \beta_h(q)\}_{q\in Q}$ are estimated using a rolling-window estimation scheme with sample size T_{is}, which we set to 200. Estimated parameters are used to compute quarter-on-quarter predictive densities up to 12 quarters ahead, from which we get S forecast paths. We then use simple approximating formulas to construct annual-average and year-on-year predictive distributions from those paths. For instance, assuming that the starting point of the forecasting exercise is the first quarter of the year, for each forecast path $s = 1, \ldots, S$, we would have for the one year-ahead annual-average transformation:

$$Z_{t+h}^{(s)} = \frac{1}{4}Y_{t-2} + \frac{2}{4}Y_{t-1} + \frac{3}{4}Y_t + Y_{t+1|t}^{(s)} + \frac{3}{4}Y_{t+2|t}^{(s)} + \frac{2}{4}Y_{t+3|t}^{(s)} + \frac{1}{4}Y_{t+4|t}^{(s)} \tag{7}$$

and for the four quarters-ahead year-on-year transformation:

$$Z_{t+h}^{(s)} = \sum_{j=1}^{h=4} Y_{t+j|t}^{(s)} \tag{8}$$

Two competing approaches are here considered. On one side, the benchmark approach, which constructs the annual-average and year-on-year forecasts by directly transforming the quarter-on-quarter forecast paths with (7) and (8), i.e., without accounting for the cross-horizon dependence. On the other, the copula approach, which in turn constructs the predictive distributions for the annual-average and yoy growth rates using first the methodology described in Section 3 to obtain joint draws of the quarter-on-quarter growth rates, and then expressions (7) and (8) to transform the joint draws.

T_{oos}, which we set to 50, denotes the out-of-sample size used to compute the correlation of the empirical PITs for the copula approach.[7]

The effective out-of-sample size available for evaluation for the annual-average (yoy) forecasts is set to 50, i.e., we simulate 200 periods of quarterly data and then produce annual-average (yoy) forecast every four quarters for horizons of one, two, and three years ahead.[8]

Results are based on 500 Monte Carlo iterations, i.e., we simulate 500 times 50 out-of-sample annual-average (yoy) predictions for horizons one, two, and three years ahead. For each of 1,...,500 Monte Carlo iterations, we evaluate the performance of the benchmark and copula approach using both relative and absolute forecasting performance measures. The relative performance measures include the log-score (LS; Amisano and Giacomini, 2007), the continous ranked probability score (CRPS; Gneiting et al., 2007) and the tick loss (Komunjer, 2013) to evaluate the predictive performance at different quantiles of the distribution. The relative performance is evaluated against the forecasts one could generate with knowledge of the true underlying DGP and the true parameter values. The absolute forecasting measures tests for the correct specification of the predictive distribution using the test of Rossi and Sekhposyan (2019), which tests for uniformity of the PIT.

4.2 Monte Carlo results

Table 1 shows rejection frequencies, computed over the 500 Monte Carlo iterations, of the null hypothesis of equal predictive ability (Panel A to B) using the unconditional test of Giacomini and White (2006) and the null hypothesis of correct specification (Panel C) evaluated using the test of Rossi and Sekhposyan (2019). The nominal size is 5%, i.e., rejection frequencies above 5% indicate that the respective approach underperforms too often than what could be expected by the nominal size. Starting with Panel A to B, the table shows that with increasing temporal

[7]To reduce the computational costs in the Monte Carlo simulations, we compute the correlation of the empirical PITs for the first annual-average forecast and use the parameter in all forecasting iterations, instead of re-estimating it in every forecasting iteration.

[8]To only produce annual-average forecasts every four quarters aims to replicate the calendar year predictions typically used in practice.

Table 1: Tests of predictive performance: rejection frequency for annual average forecast

		Normal			Skew Normal			Skew t		
θ_1	Model	h=1	h=2	h=3	h=1	h=2	h=3	h=1	h=2	h=3
		Panel A. Log-score								
0.8	Benchmark	0.59	0.74	0.67	0.47	0.68	0.62	0.47	0.69	0.64
	Copula	0.09	0.07	0.07	0.06	0.05	0.06	0.04	0.07	0.08
0.5	Benchmark	0.30	0.51	0.45	0.22	0.42	0.42	0.24	0.44	0.43
	Copula	0.05	0.06	0.09	0.04	0.10	0.09	0.03	0.07	0.09
0.1	Benchmark	0.06	0.10	0.11	0.10	0.10	0.09	0.06	0.09	0.10
	Copula	0.12	0.21	0.22	0.14	0.19	0.18	0.09	0.21	0.20
		Panel B. CRPS								
0.8	Benchmark	0.30	0.46	0.42	0.30	0.50	0.41	0.27	0.48	0.43
	Copula	0.05	0.04	0.10	0.04	0.05	0.06	0.05	0.08	0.12
0.5	Benchmark	0.15	0.29	0.27	0.14	0.25	0.24	0.17	0.27	0.26
	Copula	0.06	0.09	0.11	0.05	0.09	0.09	0.05	0.08	0.09
0.1	Benchmark	0.07	0.11	0.12	0.10	0.11	0.11	0.08	0.12	0.11
	Copula	0.11	0.18	0.19	0.12	0.15	0.14	0.09	0.17	0.15
		Panel C. PIT								
0.8	Benchmark	0.56	0.85	0.82	0.60	0.84	0.78	0.57	0.81	0.77
	Copula	0.08	0.10	0.13	0.07	0.07	0.11	0.06	0.09	0.12
0.5	Benchmark	0.29	0.46	0.47	0.28	0.41	0.40	0.36	0.46	0.46
	Copula	0.11	0.15	0.18	0.05	0.09	0.10	0.08	0.11	0.13
0.1	Benchmark	0.08	0.09	0.09	0.09	0.07	0.09	0.08	0.09	0.08
	Copula	0.13	0.20	0.20	0.08	0.10	0.12	0.08	0.14	0.12

Note: The table shows the rejection frequency of the null hypothesis of a Giacomini and White (2006) test of unconditional equal predictive ability for Panel A to B and of a test of correct specification of Rossi and Sekhposyan (2019) in Panel C. The nominal size is 5%. The equal predictive ability test in Panel A to B compares the score of the true predictive density to the score of the approach indicated by the row name. Panel C shows the results based on the Kolmogorov-Smirnov statistic. The column ρ denotes the autoregressive parameter. The column label h denotes the horizon, i.e., one-year-, two-years-, and three-years-ahead. Normal, Skew Normal, and Skew t indicate the distribution of the error terms in the DGP. Standard errors of the tests were computed using a HAC with a bandwidth $= h - 1$.

dependence, i.e. increasing ρ, the copula approach remains to have size around the nominal size whereas for the benchmark approach the null hypothesis of equal performance is rejected frequently. This reflects the fact that the copula approach takes the temporal dependence into account when constructing the annual-average predictive distributions. For very small values of ρ, the copula approach performs slightly worse than the benchmark approach, i.e., the parameter estimation uncertainty due to the copula estimation dominates the gain from taking the temporal dependence into account.

A similar picture emerges for Panel C: high temporal dependence implies that the benchmark approach is misspecified whereas the copula approach adjusts for the temporal dependence and has rejection frequencies only slightly above the nominal size.

Table 2 shows the relative forecasting performance of the benchmark and copula approach. Panel A shows the log-score of the copula approach minus the log-score of the benchmark approach, i.e., numbers above zero indicate superior performance of the copula approach. Panel B shows the CRPS of the benchmark approach relative to the copula approach, i.e., numbers smaller than one indicate a worse performance of the benchmark approach. As expected, for small serial dependence there is no improvement in the forecasting performance when using the copula approach. For medium to large values of ρ the copula approach outperforms the benchmark approach by about 3% to 7% in the case of the CRPS.

Table 3 and Table 4 show results for the year-on-year forecasts which are very similar to the annual-average results.

Table 2: Relative performance of annual average forecast

		Normal			Skew Normal			Skew t		
θ_1	Model	h=1	h=2	h=3	h=1	h=2	h=3	h=1	h=2	h=3
		Panel A. Log-score								
0.8	Copula/Benchmark	0.40	1.37	1.62	0.38	1.28	1.55	0.43	1.34	1.60
0.5	Copula/Benchmark	0.14	0.33	0.35	0.12	0.32	0.32	0.21	0.39	0.40
0.1	Copula/Benchmark	-0.01	-0.06	-0.06	-0.01	-0.03	-0.03	0.01	-0.03	-0.03
		Panel B. CRPS								
0.8	Copula/Benchmark	0.96	0.93	0.93	0.96	0.94	0.93	0.97	0.94	0.93
0.5	Copula/Benchmark	0.98	0.97	0.97	0.98	0.97	0.97	0.98	0.97	0.97
0.1	Copula/Benchmark	1.00	1.01	1.01	1.00	1.00	1.00	1.00	1.01	1.01

Note: Panel A shows the log-score of the copula approach minus the log-score of the benchmark approach, i.e., numbers above zero indicate superior performance of the copula approach. Panel B shows the CRPS of the benchmark approach relative to the copula approach, i.e., numbers smaller than one indicate a worse performance of the benchmark approach. The column ρ indicates the autoregressive parameter. The column label h denotes the horizon, i.e., one-year-, two-years-, and three-years-ahead. Normal, Skew Normal, and Skew t indicate the distribution of the error terms in the DGP.

Table 3: Rejection frequency for year-on-year forecasts

		Normal			Skew Normal			Skew t		
θ_1	Model	h=4	h=8	h=12	h=4	h=8	h=12	h=4	h=8	h=12
		Panel A. Log-score								
0.8	Benchmark	0.63	0.55	0.51	0.51	0.42	0.44	0.50	0.51	0.46
	Copula	0.10	0.07	0.06	0.08	0.05	0.06	0.06	0.07	0.08
0.5	Benchmark	0.41	0.39	0.40	0.32	0.34	0.34	0.35	0.35	0.34
	Copula	0.08	0.06	0.07	0.04	0.08	0.07	0.05	0.07	0.07
0.1	Benchmark	0.08	0.10	0.09	0.11	0.09	0.09	0.09	0.09	0.09
	Copula	0.15	0.15	0.17	0.15	0.14	0.15	0.15	0.16	0.16
		Panel B. CRPS								
0.8	Benchmark	0.36	0.33	0.33	0.36	0.32	0.28	0.31	0.33	0.32
	Copula	0.04	0.06	0.09	0.03	0.06	0.07	0.05	0.09	0.11
0.5	Benchmark	0.22	0.22	0.20	0.20	0.21	0.19	0.21	0.24	0.20
	Copula	0.07	0.09	0.09	0.05	0.08	0.08	0.07	0.07	0.08
0.1	Benchmark	0.09	0.11	0.10	0.11	0.08	0.09	0.10	0.11	0.12
	Copula	0.15	0.14	0.14	0.14	0.11	0.11	0.14	0.15	0.16
		Panel C. PIT								
0.8	Benchmark	0.68	0.68	0.67	0.64	0.64	0.62	0.58	0.61	0.58
	Copula	0.07	0.11	0.14	0.06	0.08	0.11	0.05	0.07	0.12
0.5	Benchmark	0.39	0.39	0.40	0.32	0.30	0.28	0.39	0.36	0.36
	Copula	0.11	0.14	0.14	0.05	0.07	0.08	0.10	0.11	0.10
0.1	Benchmark	0.07	0.08	0.08	0.06	0.07	0.07	0.09	0.09	0.08
	Copula	0.14	0.15	0.16	0.08	0.09	0.09	0.11	0.12	0.10

Note: The table shows the rejection frequency of the null hypothesis of a Giacomini and White (2006) test of unconditional equal predictive ability for Panel A to B and of a test of correct specification of Rossi and Sekhposyan (2019) in Panel C. The nominal size is 5%. The equal predictive ability test in Panel A to B compares the score of the true predictive density to the score of the approach indicated by the row name. Panel C shows the results based on the Kolmogorov-Smirnov statistic. The column ρ denotes the autoregressive parameter. The column label h denotes the horizon, i.e., one-year-, two-years-, and three-years-ahead. Normal, Skew Normal, and Skew t indicate the distribution of the error terms in the DGP. Standard errors of the tests were computed using a HAC with a bandwidth $= h - 1$.

Table 4: Relative performance of year-on-year forecasts

θ_1	Model	Normal			Skew Normal			Skew t		
		h=4	h=8	h=12	h=4	h=8	h=12	h=4	h=8	h=12
		Panel A. Log-score								
0.8	Copula/Benchmark	0.53	0.84	0.91	0.49	0.80	0.91	0.53	0.83	0.91
0.5	Copula/Benchmark	0.19	0.21	0.22	0.17	0.19	0.19	0.24	0.25	0.27
0.1	Copula/Benchmark	-0.03	-0.03	-0.03	-0.01	-0.01	-0.01	-0.00	-0.02	-0.01
		Panel B. CRPS								
0.8	Copula/Benchmark	0.96	0.95	0.94	0.96	0.95	0.95	0.96	0.95	0.95
0.5	Copula/Benchmark	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
0.1	Copula/Benchmark	1.01	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00

Note: Panel A shows the log-score of the copula approach minus the log-score of the benchmark approach, i.e., numbers above zero indicate superior performance of the copula approach. Panel B shows the CRPS of the benchmark approach relative to the copula approach, i.e., numbers smaller than one indicate a worse performance of the benchmark approach. The column ρ indicates the autoregressive parameter. The column label h denotes the horizon, i.e., one-year-, two-years-, and three-years-ahead. Normal, Skew Normal, and Skew t indicate the distribution of the error terms in the DGP.

5 Empirical applications

5.1 Large bivariate exercise

In this section, we provide the results of a large-scale forecasting exercise based on monthly data from FRED-MD (McCracken and Ng, 2016). In the construction of the forecasting environment, we closely follow McCracken and McGillicuddy (2019) and consider a bivariate system, $Z_t = (Y_t, X_t)'$, consisting of a pair of two stationary series randomly drawn from the series available in FRED-MD.[9] We first compute density forecasts for month-on-month values y_{t+h}, with $h = 1, \ldots, 12$ months, and then we use these predictive densities to compute quarter-on-quarter density forecasts through our proposed copula approach. Density forecasts are obtained through an autoregressive distributed lag (ARDL) DMS regression, estimated via OLS at each forecast origin:

$$y_{t+h|t} = \alpha + \sum_{j=0}^{p-1} \beta_j y_{t-j} + \sum_{j=0}^{p-1} \gamma_j x_{t-j} + \varepsilon_{t+h} \tag{9}$$

For the sake of simplicity, we only consider the first month of each quarter as a forecast origin. The number of lags p is either fixed at four or selected through the Bayesian Information Criterion (BIC) among $p \in \{0, 1, \ldots, 12\}$.

We consider two samples: a full sample starting in 1959:M1 and a reduced sample starting in 1984:M1, covering only the Great Moderation period. The sample ends in both cases in 2019:M12. The sample is partitioned into a forecasting part and an in-sample part for estimation. The forecasting sample is further partitioned into two blocks: the first block ranges from 2003:Mh to 2010:M12, in which we compute the *initial* set of PITs that we use to estimate the copula parameters; the second block ranges from 2010:Mh to 2019:M12, and it is used for (pseudo)[10]

[9] Stationarity is ensured by taking first- or second-order differences or log-differences, depending on the series.
[10] We do not use real-time data vintages.

out-of-sample evaluation. Estimation is carried out recursively using a rolling-window sample, updating at each forecast origin the model parameters, the PITs, and the copula parameters.

As a datasource we use the February 2023 vintage of FRED-MD. After dropping series not meeting some minimal conditions[11] , the number of variables used in the bivariate exercise is 101, organized into 5 different groups as in Marcellino et al. (2006) and McCracken and McGillicuddy (2019): (1) income, output, sales, and capacity utilization; (2) employment and unemployment; (3) construction, inventories, and orders; (4) interest rates and asset prices; (5) nominal prices, wages, and money.

Then, 500 random pairs of y and x are selected from the database such that y and x come from distinct groups and an equal number of series pairs (y, x) comes from each of the 10 possible group pairings. For each permutation of the series, we compute qoq density forecasts at horizon $\widetilde{h} = 1, \ldots, 4$ from the original mom densities using either the copula approach or the benchmark approach, and we then compare their forecasting performance using their average log-score, CRPS, and tail quantile scores (QS 10% and QS 90%) over the out-of-sample. For each lag selection method and forecast horizon $\tilde{h}$, we hence have a distribution of 1,000 average score ratios (or score differentials in the case of the log-score), for which we compute the median score value. To grasp the statistical significance of these results, we also compute a test of equal (unconditional) predictive accuracy (Giacomini and White, 2006). We provide rejection rates at 5% level of the null hypothesis that the copula approach outperforms the benchmark approach, i.e., we perform one-sided tests testing for the alternative that the copula approach provides better density forecasts than the benchmark approach.

Results are reported in Table 5 and show that the copula approach outperforms the benchmark approach for several quarter ahead forecasts, independently of the lag length and the specific scoring rule considered. In other words, for most target variable and predictor combination, the copula approach delivers better forecasting results than the benchmark approach.

[11]We exclude from the dataset three series starting after the 1970 (new orders for consumers goods, new orders for non-defense capital goods, trade weighted US dollar index), one series presenting missing values (consumer sentiment index), and one series switching to negative over the sample (non-borrowed reserves of depository institutions). In addition, we exclude 21 series that should be used in levels or log-levels, as we focus on series that can be expressed in first/second difference (or log-difference) in our application.

Table 5: Relative performance of copula approach for quarter-on-quarter forecasts

Lag length	Statistics	Great moderation				Full sample			
		$\widetilde{h}=1$	$\widetilde{h}=2$	$\widetilde{h}=3$	$\widetilde{h}=4$	$\widetilde{h}=1$	$\widetilde{h}=2$	$\widetilde{h}=3$	$\widetilde{h}=4$
		Panel A. Log-score							
AR(4)	Median	-0.01	0.00	0.17	0.26	-0.01	0.00	0.19	0.30
	Test 1S	0.05	0.27	0.64	0.67	0.05	0.28	0.65	0.68
	Test 2S	0.26	0.40	0.75	0.78	0.28	0.45	0.77	0.82
BIC	Median	-0.02	0.00	0.10	0.20	-0.01	-0.01	0.14	0.28
	Test 1S	0.04	0.28	0.56	0.65	0.04	0.26	0.59	0.69
	Test 2S	0.23	0.46	0.70	0.77	0.25	0.46	0.72	0.78
		Panel B. CRPS							
AR(4)	Median	1.00	1.01	1.10	1.21	1.00	1.01	1.12	1.23
	Test 1S	0.08	0.25	0.62	0.73	0.08	0.27	0.66	0.74
	Test 2S	0.29	0.42	0.74	0.78	0.30	0.47	0.78	0.79
BIC	Median	1.00	1.00	1.08	1.17	1.00	1.00	1.09	1.20
	Test 1S	0.06	0.24	0.52	0.69	0.06	0.23	0.57	0.72
	Test 2S	0.24	0.45	0.74	0.78	0.26	0.47	0.70	0.77
		Panel C. QS 10%							
AR(4)	Median	0.97	0.99	1.27	1.56	0.98	0.99	1.28	1.57
	Test 1S	0.05	0.24	0.75	0.85	0.04	0.23	0.77	0.87
	Test 2S	0.28	0.56	0.77	0.85	0.26	0.58	0.77	0.87
BIC	Median	0.97	0.99	1.21	1.42	0.98	0.98	1.24	1.48
	Test 1S	0.04	0.22	0.64	0.82	0.03	0.22	0.69	0.85
	Test 2S	0.26	0.54	0.73	0.84	0.25	0.57	0.74	0.86
		Panel D. QS 90%							
AR(4)	Median	0.99	1.03	1.27	1.48	0.99	1.04	1.27	1.52
	Test 1S	0.05	0.34	0.59	0.66	0.03	0.35	0.63	0.69
	Test 2S	0.21	0.39	0.70	0.73	0.19	0.43	0.74	0.75
BIC	Median	1.00	1.03	1.18	1.39	0.99	1.03	1.20	1.46
	Test 1S	0.07	0.32	0.57	0.67	0.04	0.32	0.59	0.68
	Test 2S	0.21	0.39	0.70	0.76	0.19	0.41	0.72	0.76

Note: Panel A shows the log-score of the copula approach minus the log-score of the benchmark approach, i.e., numbers above zero indicate superior performance of the copula approach. Panels B, C, and D show respectively the CRPS, the QS10%, and the QS90% of the benchmark approach relative to the copula approach, i.e., numbers larger than one indicate a worse performance of the benchmark approach. The column "Lag length" indicates the lag of the underlying ARDL regressions(fixed $p = 4$ or BIC selection). The column label "h" denotes the horizon at quarterly frequency, i.e., one-quarter-, two-quarters-, three-quarters, and four-quarters-ahead.

5.2 Inflation-at-Risk

In this section, we provide estimates of inflation at risk for year-on-year and annual-average inflation based on the U.S. Consumer Price Index. Consider the following situation: an institution has predictive densities for year-on-year inflation, which is the benchmark inflation rate for central banks such as the Federal Reserve or the European Central Bank (ECB). However, while the year-on-year inflation rate is a benchmark, to get a more complete picture of the inflation environment, the institutions tend to extend their set of forecasts to include annual-average predictive densities; note that, in fact, both the ECB and the Federal Reserve publish annual-average forecasts as part of their institutional projection exercises. Importantly, the year-on-year and annual-average predictions need to be coherent, i.e., they should be based on the same predictors, model type, and the central tendency of the forecasts across the two frequencies should be very similar.

To emulate this situation, we use a quantile regression model to produce quantile predictions for monthly year-on-year inflation for up to 12 months ahead. We then use the copula approach to combine the monthly year-on-year inflation predictive distributions into annual-average predictive distributions.

The underlying price index is the monthly and seasonally adjusted Consumper Price Index for all Urban Consumers[12] (henceforth only CPI) from 1960 to 2022. The baseline model predicts year-on-year inflation, computed via the log-difference of t and $t-12$.

The forecasts of the monthly year-on-year inflation rate are produced via a quantile regression model that uses a Lasso to select among a number of potential predictors. Estimating a quantile regression model with L1 penalization amounts to solving the following objective function for β:

$$\min_{\beta \in \mathbb{R}^p} \sum_{t=1}^{T} \rho_\tau(y_{t+h} - x_t'\beta) + \frac{\lambda\sqrt{u(1-u)}}{n} \sum_{j=1}^{p} \widehat{\sigma}_j^2 |\beta_j|, \tag{10}$$

where y_{t+h} denotes the monthly year-on-year inflation rate, $\rho_\tau(z) = (\tau - \mathbb{1}\{t \leq 0\})z$ denotes the tick function, λ is a hyperparameter that determines the degree of penalization, p is the number of predictors, $h = 1, ..., 12$ is the forecast horizon, and $\widehat{\sigma}_j^2 = \sum_{t=1}^{T} x_{i,t}^2$. We are estimating the model using the code made available by Belloni and Chernozhukov (2011) as well as their calibration for λ.

The predictor vector x_t contains a maximum of 22 predictors, similar to the predictors used in Korobilis (2017), y_t and y_{t-1}[13] and an intercept; see Appendix A for details on the data series. Note that because not all series cover the entire range of the data, not all predictors enter the model in all forecast origins. The predictions generated by eq. (10) are then combined into annual-average inflation predictions via the copula and benchmark approach, as described below.

We use the model in eq. (10) to produce (pseudo)[14] out-of-sample forecasts for forecast horizons of $h = 1, ..., 12$ months ahead, with the first forecast origin being 1974:M12 and the last forecast origin being 2021:M12 for $h = 1, ..., 12$. At each forecast origin, the quantile regression is re-estimated over a rolling window of 15 years of monthly data. For instance, for the forecast origin 1974:M12, the first observation used in the model estimation is 1960:M1.

[12] Results are robust to using the non-seasonally adjusted CPI for all items, which has the mnemonic USACPIALLMINMEI on FRED.

[13] Results are robust to increasing the number of lags to six.

[14] The forecasts are pseudo out-of-sample since we do not use real-time data. Henceforth, we refer to the forecasts simply as "out-of-sample".

The models produce quantile predictions for $\tau \in [0.01, 0.02, ..., 0.98, 0.99]$[15] and we calculate the predictive cdf based on linearly interpolating between adjacent quantiles. Details are provided in Appendix A.

To construct annual-average forecasts based on the predictions of eq. (10), we then use 10 years[16] of data to estimate the copula parameter over a rolling window starting with the years 1975:M1 to 1984:M12. To be precise, we start by evaluating the empirical PITs of the predictive distributions of monthly year-on-year inflation from 1975:Mh to 1984:Mh for horizon $h = 1, ..., 12$. We then estimate the copula parameter based on the empirical PITs, and use the estimated copula parameters in combination with the forecasts for 1985:Mh, $h = 1, ..., 12$, to construct an annual-average inflation predictive distribution for 1986. Next, we use the predictive distributions for year-on-year inflation from 1976:Mh to 1986:Mh for horizon $h = 1, ..., 12$, and re-do the steps to construct the annual-average inflation predictive distribution for 1987. We then repeat this until we have an annual-average forecast for each year from 1986 to 2022. In addition, we construct annual-average forecasts as described above but using the benchmark approach, which is identical to setting all off-diagonal elements of R equal to zero.

Therefore, the out-of-sample evaluation period for annual-average inflation forecasts is from 1986 to 2022, which leaves us with a sample of 37 out-of-sample (calendar year) annual-average predictive distributions.

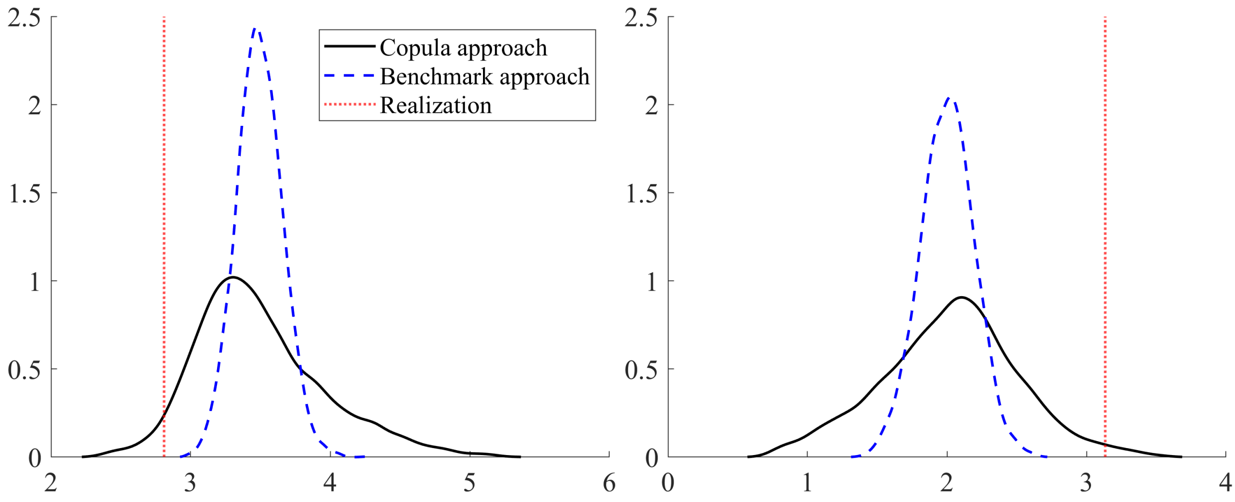
The ratio of the CRPS of the copula and the benchmark approach is 0.90, i.e., the copula approach delivers a 10% better performance. The better performance of the copula approach relative to the benchmark approach is statistically significant at the 1% level when doing a forecast comparison test using Giacomini and White (2006).

The ratio of the tick loss function for the 10% and 90% quantile is 0.70 and 0.88, i.e., the copula approach delivers a 30% and 10% better performance for predicting the risk of low inflation and the risk of high inflation, respectively. The better performance of the copula approach relative to the benchmark approach is statistically significant at the 1% level for the 10% quantile when doing a forecast comparison test using Giacomini and White (2006).

Figure 2 illustrates the superior ability of the copula approach to predict tail-risk. Panel (a) shows the predictive density for the annual-average inflation for 2001 and panel (b) shows the predictive density for the annual-average inflation for 2011. In both cases the realizations receive an ex-ante probability of close to zero by the benchmark approach whereas the copula-based annual-average inflation forecast density has fatter tails due to the high correlation between adjacent months' year-on-year growth rates. Note that the distribution of the copula-based approach exhibits notably stronger asymmetry.

[15]Quantile crossing is dealt with by sorting the predictive quantiles.
[16]Results are robust to using five or 15 years as training data for the copula approach.

Figure 2: Annual-average predictive densities for inflation



(a) For 2001 (b) For 2011

Note: The figure shows the results for the annual-average inflation predictive density for 2001, panel (a), and for 2011, panel (b), for both the copula approach (solid line) as well as the benchmark approach (dashed line), alongside the realization (dotted line).

5.3 Growth-at-Risk

In a recent influential contribution, Adrian et al. (2019) use quantile regressions to produce *direct* forecasts of quarterly quarter-on-quarter U.S. real GDP growth and show that financial conditions, captured by the National Financial Conditions Index (NFCI), are an important predictor for downside risk to GDP growth. Their specification produces forecasts for quarterly quarter-on-quarter growth rates and we apply our methodology to transform their quarter-on-quarter forecasts into annual-average GDP growth predictive densities. This allows us to use the original predictive densities of Adrian et al. (2019) as the basis for our annual-average predictive distributions.

The total data sample of Adrian et al. (2019) ranges from 1973:Q1 to 2015:Q4 and starting with forecast origin 1993:Q1 the authors produce (pseudo) out-of-sample forecasts using quantile regressions for up to four quarters ahead.[17] Thus using their exact specification, this leaves us with a series of out-of-sample forecasts from 1993:Q(1+h) to 2015:Q4, for forecast horizons h=1,...,4.

To construct annual-average predictive distributions, we proceed as follows. We start with evaluating the series of predictive distributions at forecast targets 1993:Q(1+h) to 2001:Q(1+h-1), h=1,...,4, at their realizations to obtain a series of 32 empirical PITs for each forecast horizon h=1,...,4. Given the series of empirical PITs, we calculate the copula parameter as described in Section 3. The resulting estimate of the copula parameter together with the quarter-on-quarter predictive distributions for horizon h=1,...,4, with origin 2001:Q4 (and forecasting target 2002:Q1, to 2002:Q2, 2002:Q3, and 2002:Q4), then allow us to compute an annual-average growth predictive distribution for 2002. This annual-average growth predictive distribution is based on the original quarterly predictive distributions of Adrian et al. (2019) and takes into account the

[17] The predictors of their preferred specification are the own lag of real GDP growth and the NFCI. For details on the exact specification and variable definitions, please see Adrian et al. (2019).

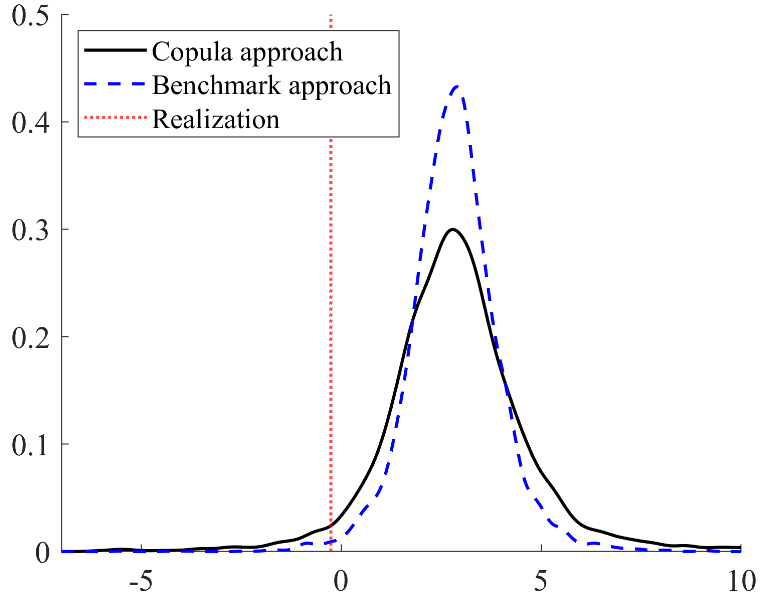
serial correlation across the quarterly growth rates.

Then, we move four quarters ahead and repeat the exercise to construct the annual-average predictive distribution for 2003. In total, we repeat this until we have a series of 14 annual-average predictive distributions for 2002 to 2015.

In addition, we construct the benchmark annual-average predictive distributions which are also based on the Adrian et al. (2019) quarterly quarter-on-quarter predictive distributions, but ignore the serial correlation between the quarterly growth rates.

Figure 3 shows the results for the annual-average predictive density of 2008 for both the copula-approach (solid line) as well as the benchmark approach (dashed line), alongside the realized annual-average growth rate in 2008 (dotted line). The copula-approach leads to annual-average forecasts with larger tails due to the positive correlation between the quarterly growth rates; the rank correlation, i.e., the elements in the Gaussian copula correlation matrix, is around 50% to 60% for adjacent quarters, depending on the horizon. Importantly, the larger tails of the copula-approach help to assess the downside risk for the annual-average forecast of 2008, i.e., during the onset of the financial crisis. This underlines the usefulness of our approach and provides anecdotal evidence that risk assessments can be misleading if the serial correlation is not taken into account for the transformation of predictive distributions to different frequencies.

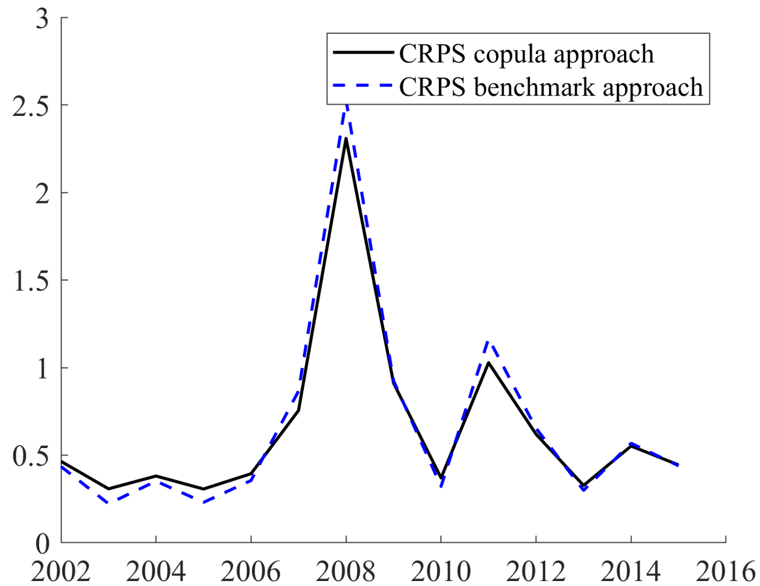
Figure 3: Annual-average predictive densities for 2008



Note: The figure shows the results for the annual-average predictive density of 2008 for both the copula approach (solid line) as well as the benchmark approach (dashed line), alongside the realization (dotted line).

Figure 4 shows the negative of the continuous ranked probability score (CRPS; Gneiting et al., 2007) of the copula approach (solid line) and benchmark approach (dashed line) for the 14 annual-average predictive distributions. Importantly, larger values imply a worse forecasting performance. While the sample is too short for formal testing procedures, the overall ratio of the CRPSs of the two approaches points to a 2% gain of the copula approach over the benchmark approach in terms of their CRPS score.

Figure 4: CRPS values of annual-average forecasts



Note: The figure shows the negative of the CRPS of the copula approach (solid line) and benchmark approach (dashed line) for the 14 annual-average predictive distributions. Importantly, larger values imply a worse forecasting performance.

6 Conclusion

In this work, we provide a method to combine *direct* forecasts to obtain new predictive objects that are function of several horizons. Our methdology is useful in a situation where the forecaster has a set of *direct* forecasts available and has to use the same set of *direct* predictive densities to construct a new predictive object; for instance, the model produces *direct* quarter-on-quarter growth rates and the forecasters also needs annual-average predictions. These type of situations typically arise, but are not limited to, in institutions where the forecasting specification is rigid.

In a Monte Carlo exercise, we show that our methodolgy outperforms the benchmark approach whenever the serial correlation across different forecasting horizons is not extremely low. In the relative forecasting exercise, the copula approach delivers significantly better results than the benchmark approach. In the absolute forecasting exercise the copula approach provides density forecasts that pass a correct specification test based on evaluating the uniformity of the PIT, whereas the benchmark approach fails to pass this test whenever the serial correlation is at least moderately high.

In the first empirical application, we transform the quarter-on-quarter *direct* forecasts of U.S. real GDP growth of Adrian et al. (2019) into annual-average forecasts, and provide anecdotal evidence that the copula approach provides a better forecasts of the growth at risk during the great recession period.

In the second empirical application, we show the usefulness of the our approach of transforming year-on-year predictive densities for inflation into annual-average predictive densities. The copula approach significantly outperforms the benchmark approach both in terms of the CRPS and quantile tick loss evaluation.

In a third empirical application, we investigate in a large-scale forecasting exercise, based on monthly data from FRED-MD (McCracken and Ng, 2016), the performance of our methodology for a large number of outcome variable and predictor combinations. In this exercise, we transform month-on-month predictive densities to quarter-on-quarter density forecasts through our proposed copula approach and results show that copula approach outperforms the benchmark approach for the majority of outcome variable and predictor combinations.

References

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

Amisano, G. and Giacomini, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics*, 25(2):177–190.

Azzalini, A. and Capitanio, A. (2003). Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t-Distribution. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 65(2):367–389.

Belloni, A. and Chernozhukov, V. (2011). L1-penalized quantile regression in high-dimensional sparse models. *The Annals of Statistics*, 39(1):82–130.

Ferrara, L., Mogliani, M., and Sahuc, J.-G. (2022). High-frequency monitoring of growth at risk. *International Journal of Forecasting*, 38(2):582–595.

Giacomini, R. and White, H. (2006). Tests of Conditional Predictive Ability. *Econometrica*, 74(6):1545–1578.

Gneiting, T., Balabdaoui, F., and Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2):243–268.

Komunjer, I. (2013). Chapter 17 - quantile prediction. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2 of *Handbook of Economic Forecasting*, pages 961–994. Elsevier.

Korobilis, D. (2017). Quantile regression forecasts of inflation under model uncertainty. *International Journal of Forecasting*, 33(1):11–20.

Marcellino, M., Stock, J. H., and Watson, M. W. (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics*, 135(1-2):499–526.

Mariano, R. S. and Murasawa, Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, 18(4):427–443.

McCracken, M. W. and McGillicuddy, J. T. (2019). An empirical investigation of direct and iterated multistep conditional forecasts. *Journal of Applied Econometrics*, 34(2):181–204.

McCracken, M. W. and Ng, S. (2016). FRED-MD: A Monthly Database for Macroeconomic Research. *Journal of Business & Economic Statistics*, 34(4):574–589.

Rossi, B. and Sekhposyan, T. (2019). Alternative tests for correct specification of conditional predictive densities. *Journal of Econometrics*, 208(2):638–657.

Sklar, A. (1959). Fonctions de repartition a n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris (in French)*, 8:229–231.

Appendix

Appendix A Inflation at Risk

A.1 Data

Table A.1 shows the predictors used in Section 5.2. All data is seasonally adjusted were applicable. The transformation codes imply the following: 1 — no transformation; 4 — $\log(x_t)$; 5 — $100[\log(x_t)$ - $\log(x_{t-12})]$

Table A.1: Predictors for inflation at risk forecasts

Variable	Transformation	Mnemonic
Aggregate weekly hours	5	AWHI
Commercial + industrial loans	5	BUSLOANS
Labor force participation rate	4	CIVPART
Consumer loans	5	CONSUMER
CPI All Urban Consumers	5	CPIAUCSL
Canadian dollar to U.S. exchange rate	5	EXCAUS
Japanese Yen to U.S. exchange rate	5	EXJPUS
British Pound to U.S. exchange rate	5	EXUSUK
Federal Funds Target	1	FEDFUNDS
Private Housing starts	4	HOUST
New Family houses sold	4	HSN1F
Industrial production	5	INDPRO
Fixed-rate 30-year mortgage rate	1	MORTG
Bank prime loan rate	1	MPRIME
Motor Vehicle assemblies	1	MVATOTASSS
Total non-farm employees	5	PAYEMS
Real estate loans	5	REALLN
Capacity utilization	4	TCU
Number unemployed for 15 weeks & over	4	UEMP15OV
Number unemployed for less than 5 weeks	4	UEMPLT5
University of Michigan: consumer sentiment	1	UMCSENT
Unemployment rate	1	UNRATE
WTI spot price	5	WTISPLC

Note: All data was downloaded from the database FRED of the St. Louis Federal Reserve Bank.

A.2 Quantile interpolation

Let $F_{t+h|t}(y_{t+h})$ denotes the predictive cumulative distribution function in t and h horizons ahead, evaluated at y_{t+h}. Then, given the set $\{Q_{t+h|t}(\tau_i)\}_{i=1}^{99}$ of predictive quantiles, we compute $F_{t+h|t}(y_{t+h})$ as follows:

$$F_{t+h|t}(y_{t+h}) = \tau_i + \frac{\tau_{i+1} - \tau_i}{Q_{t+h|t}(\tau_{i+1}) - Q_{t+h|t}(\tau_i)}(y_{t+h} - Q_{t+h|t}(\tau_i)), \tag{11}$$

where $Q_{t+h|t}$ denotes predictive value of quantile τ_i, and τ_i and τ_{i+1} are such that $y_{t+h} \in [Q_{t+h|t}(\tau_i), Q_{t+h|t}(\tau_{i+1})]$. For values of $y_{t+h} < Q_{t+h|t}(\tau_1)$ and $y_{t+h} > Q_{t+h|t}(\tau_{99})$, we approximate the slope as $\frac{\tau_2 - \tau_1}{Q_{t+h|t}(\tau_2) - Q_{t+h|t}(\tau_1)}$ and $\frac{\tau_{99} - \tau_{98}}{Q_{t+h|t}(\tau_{99}) - Q_{t+h|t}(\tau_{98})}$ and the distance as $(Q_{t+h|t}(\tau_1) - y_{t+h})$ and $(y_{t+h} - Q_{t+h|t}(\tau_{99}))$.

Similarly, we sample from the distribution given by the conditional quantiles $Q_{t+h|t}(\tau_i)$ as follows. Let u_j denote the j-th draw from the uniform distribution and $u_j \in [\tau_i^j, \tau_{i+1}^j]$. Then,

$$y_{t+h|t}^j = Q_{t+h|t}(\tau_i^j) + \frac{Q_{t+h|t}(\tau_{i+1}^j) - Q_{t+h|t,}(\tau_i^j)}{\tau_{i+1}^j - \tau_i^j}(u_j - \tau_i^j), \tag{12}$$

where $y_{t+h|t}^j$ denotes draw j of the predictive distribution for y_{t+h}, conditional on information in t. The two endpoints are treated analogously to the procedure described for eq. (11).

A.3 Annual-average regression

This section illustrates, using a simplified example, why designing the forecasting regressions in an annual-average format does not lead to a superior forecasting performance. Assume the DGP at the quarterly frequency takes the following form:

$$y_t = \rho y_{t-1} + e_t \tag{13}$$

with $|\rho| < 1$ and $e_t \sim N(0, \sigma^2)$. The aim is to predict, the annual-average, simplistically defined as $X_t = \frac{1}{2}(y_t + y_{t-1})$. This simplification implies that a calendar year consists of two periods, the respective calendar year annual-average growth rates are $X_t, X_{t-2}, X_{t-4}, ...$, and we refer to the time periods denoted by t as quarters, i.e., a calendar year consists of two quarters.[18] The annual-average can also be written as a process in the "calendar-year" annual-averages:

$$X_t = \frac{1}{2}(y_t + y_{t-1}) \tag{14}$$

$$= \rho^2 \underbrace{\frac{1}{2}(y_{t-2} + y_{t-3})}_{X_{t-2}} + \underbrace{\frac{1}{2}(e_t + (1+\rho)e_{t-1} + \rho e_{t-2})}_{u_t} \tag{15}$$

$$= \rho^2 X_{t-2} + u_t \tag{16}$$

Consider now a situation where the latest available observation is t and that the aim is to produce predictions of X_{t+2} and X_{t+4}, i.e., the next and the subsequent calendar year.

[18] This timing convention and that the annual-average X_t is the simple mean of y_t and y_{t-1} is made to keep the analytical derivations simple.

First, the "annual-average" regression suggested by eq. (16) suggest a forecast of $\hat{X}_{t+2,AA} = \rho^2 X_t$.[19] The expected squared forecast error is then

$$\text{SF}_{t+2,AA} = E[(X_{t+2} - \hat{X}_{t+2,AA})^2] = E[(\frac{1}{2}(e_{t+2} + (1+\rho)e_{t+1} + \rho e_t)^2] \tag{17}$$

$$= \frac{(1+\rho)^2\sigma^2}{4} + \frac{\rho^2\sigma^2}{4} + \frac{\sigma^2}{4} \tag{18}$$

Second, the higher-frequency approach would construct the forecast as follows: $\hat{X}_{t+2,q} = \frac{1}{2}(\hat{y}_{t+2} + \hat{y}_{t+1})$, where $\hat{y}_{t+2} = \rho^2 y_t$ and $\hat{y}_{t+1} = \rho y_t$. The expected squared forecast error is

$$\text{SF}_{t+2,q} = E[(X_{t+2} - \hat{X}_{t+2,q})^2] = E[(\frac{1}{2}(y_{t+2} + y_{t+1}) - \frac{1}{2}(\rho^2 y_t + \rho y_t))^2] \tag{19}$$

$$= \frac{1}{4}E[(y_{t+2} - \rho^2 y_t + y_{t+1} - \rho y_t)^2] \tag{20}$$

$$= \frac{1}{4}E[(e_{t+2} + \rho e_{t+1} + e_{t+1})^2] \tag{21}$$

$$= \frac{\sigma^2}{4} + \frac{(1+\rho)^2\sigma^2}{4}. \tag{22}$$

The difference between $\text{SF}_{t+2,AA}$ and $\text{SF}_{t+2,q}$ is equal to $\frac{\rho^2\sigma^2}{4}$, such that the expected mean forecast of the annual-average regression is worse than that of the quarter-on-quarter regression whenever $\rho \neq 0$. The term $\frac{\rho^2\sigma^2}{4}$ is a consequence of the annual-average regression in eq. (15) having the additional term ρe_{t-2}. The intuition that the quarter-on-quarter regressions have a better forecasting performance then the annual-average regression is then that the annual-average regression assigns an equal weight to the latest and prior-to-latest observation, which results in an inefficient use of information.

The two calendar year ahead regressions have the following expected squared forecast errors.

$$\text{SF}_{t+4,AA} = E[(X_{t+4} - \hat{X}_{t+4,A})^2] = E[(u_{t+4} + u_{t+2})^2] \tag{23}$$

$$= \frac{1}{4}(\sigma^2 + (1+\rho)^2\sigma^2 + (\rho+\rho^2)^2\sigma^2 + (\rho^2+\rho^3)^2\sigma^2 + \rho^6\sigma^2) \tag{24}$$

and

$$\text{SF}_{t+4,q} = E[(X_{t+4} - \hat{X}_{t+4,q})^2] = \frac{1}{4}(\sigma^2 + (1+\rho)^2\sigma^2 + (\rho+\rho^2)^2\sigma^2 + (\rho^2+\rho^3)^2\sigma^2). \tag{25}$$

The difference between $\text{SF}_{t+4,AA}$ and $\text{SF}_{t+4,q}$ is $\rho^6\sigma^2$. This implies that for any value of $\rho \neq 0$ and $|\rho| < 1$, the quarter-on-quarter regression performs better than the annual-average regressions. At the same time, the relative gain has become smaller with the forecasting horizon.

[19] Throughout the section we abstract from parameter estimation error and ignore here the endogeneity issue in estimating ρ^2 in the annual-average specification because X_t and u_{+2} are correlated.