BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working Papers)

The structural Theta method and its predictive performance in the M4-Competition

by Giacomo Sbrana and Andrea Silvestrini

June 2024

Number 1457

BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working Papers)

The structural Theta method and its predictive performance in the M4-Competition

by Giacomo Sbrana and Andrea Silvestrini

Number 1457 - June 2024

The papers published in the Temi di discussione *series describe preliminary results and are made available to the public to encourage discussion and elicit comments.*

The views expressed in the articles are those of the authors and do not involve the responsibility of the Bank.

ISSN 2281-3950 (online)

Designed by the Printing and Publishing Division of the Bank of Italy

THE STRUCTURAL THETA METHOD AND ITS PREDICTIVE PERFORMANCE IN THE M4-COMPETITION

by Giacomo Sbrana* and Andrea Silvestrini**

Abstract

The Theta method is a well-established prediction benchmark widely used in forecast competitions. Introduced more than 20 years ago, this method has received significant attention, with several authors proposing different variants to improve its performance. This paper considers the multiple sources of error version of the Theta model, belonging to the family of structural time series models, and investigates its out-of-sample forecast performance using the extensive M4-Competition dataset, which includes 100,000 time series. We compare the proposed structural Theta model against several benchmarks, including all variants of the Theta method. The results clearly demonstrate its remarkable predictive abilities as it outperforms all its variants and competitors, emerging as a solid benchmark for use in forecast competitions.

JEL Classification: C13, C22.

Keywords: Theta method, state-space models, Kalman filter, M4-Competition, predictive accuracy.

DOI: 10.32057/0.TD.2024.1457

* Giacomo Sbrana, NEOMA Business School. E-mail: giacomo.sbrana@neoma-bs.fr.

** Andrea Silvestrini, Bank of Italy, DG for Economics, Statistics and Research.
E-mail: andrea.silvestrini@bancaditalia.it.

1. Introduction[1]

The Theta method was introduced by [2] as a univariate forecasting algorithm particularly useful for generating predictions of time-series data. It is based on the idea of decomposing a time series into a trend component, representing the long-term pattern of the data, and a residual one, capturing the short-term random fluctuations around the trend. In order to do so, the original Theta method adjusts the local curvature of the series by applying two coefficients, called theta coefficients, to the second differences of the data. The theta coefficients control whether the local curvature is decreased (deflating the series) or increased (dilating the series), thereby amplifying either the long-term or short-term component, respectively [24]. The resulting two distinct components, referred to as "theta lines", approximate the long and short-term dynamics of the time series, which are subsequently projected and then combined for forecast calculation. This method, as originally proposed by [2], is often referred to as the "classic" Theta method to distinguish it from other variants and subsequent extensions.

Forecasts produced by the classic Theta method are equivalent to those delivered by a simple exponential smoothing with drift, with the restriction that the drift parameter is set to half the slope of the linear trend fitted to the data, see [14]. However, as clearly stated by [24], pp. 11-12, in a standard simple exponential smoothing with drift both smoothing parameters are typically either set or optimized simultaneously in the original time series. As a consequence, regardless of the optimization method used, the smoothing parameters coincide with those used in the classic Theta model only by coincidence. For this reason the forecasts generated by the two models are generally different, despite the functional form of the model being the same.

Extensions of the classic Theta method have been put forward by [31], who proposed

[1] The views expressed herein are those of the authors and do not necessarily reflect those of the Bank of Italy or the Eurosystem. The authors are grateful to Antonio Di Cesare and Raffaela Giordano, two anonymous reviewers, Fabio Busetti, Davide Delle Monache and Ivan Petrella for useful comments and suggestions on a previous draft.

a multivariate version of the original univariate specification, and by [9], who introduced the Dynamic Optimised Theta Model (DOTM), a state-space formulation that selects the best short-term theta line and revises dynamically the long-term theta line. More recently, [29] suggested three further extensions on the Theta's original framework, namely, considering both linear and non-linear trends, allowing the slope of the trend to be adjusted, and introducing a new multiplicative representation along with the standard additive one.

The classic Theta has been applied in various contexts and exhibited remarkable performance in the M3 competition [18].[2] As such, it is nowadays considered a solid forecast method to beat. The Theta method is also extensively employed by large multinational organisations (e.g., Uber, Amazon and Bosch) in the private sector. For example, Uber leverages quantitative forecasting methods to accurately predict user supply and demand. Specifically, among statistical models, Uber utilizes the Theta approach, known for its computational efficiency and demonstrated success in Uber's time series data.[3] Similarly, Amazon Forecast uses the Theta method within the Exponential Smoothing (ETS) built-in algorithm.[4] Additionally, companies like Bosch have adopted this method for predicting demand and sales due to its proven performance and fast implementation. However, as argued by [28], despite its good empirical performance limited academic research has been conducted to generalize its reach and exploit its full potential. This paper aims to address this issue by introducing the structural Theta model, or Multiple Source of Error (MSOE) Theta model, and investigating its prediction performance. The widespread application of the Theta method by leading companies in the private sector clearly underscores the impact of our research.

The structural Theta model examined in this work is an observationally-equivalent variant of the single source of error (SSOE) simple exponential smoothing with drift [14]. Unlike this latter model, the MSOE version incorporates distinct errors that affect

[2]Specifically, it was 3.8% more accurate than the Comb method, see [19].

[3]`https://www.uber.com/en-GB/blog/forecasting-introduction/`

[4]`https://docs.aws.amazon.com/forecast/latest/dg/aws-forecast-recipe-ets.html`

both the states and observations in its state-space formulation [3]. In contrast, the simple exponential smoothing with drift relies on a single perturbation term, which corresponds to the one-step-ahead forecast error, to drive the entire system [14].

The structural Theta model can also be connected to the extensive body of literature studying the separation of time series into a long-term trend component and a medium-term business cycle, bearing in mind that the random walk with a drift has often been used to estimate the trend (e.g., Beveridge-Nelson decomposition), see [25]. This literature is closely linked to the broader debate in macroeconomics regarding the significance of permanent and transitory factors in explaining GDP growth ([32]; [22]).

It is noteworthy that the predictive accuracy of neither the structural Theta model nor the SSOE simple exponential smoothing with drift has been assessed within the framework of the M4 forecast competition, despite the fact that the classic Theta has been employed as a benchmark method. This paper effectively fills this gap by specifically evaluating the forecast accuracy of the structural Theta model using the M4-Competition [20], which extends the results of the previous three Makridakis Competitions. The M4-Competition provides a comprehensive database that encompasses an extensive collection of time-series data from various industries and economic sectors, spanning different time frequencies. By including a large number of time series and facilitating the calculation of forecasts, the M4-Competition plays a crucial role in enabling more reliable and robust assessments of forecasting methods in an empirical context.

We decided to focus on the structural Theta model after a close look at the properties of the M4 series. Our analysis reveals that, for the vast majority of them, standard unit root tests do not reject the presence of a unit root. Moreover, among the integrated series, the presence of a drift is detected in more than one-third of them. These two factors are the crucial ingredients underlying the dynamic properties of the random walk with drift plus noise model, being the state-space representation of our proposed approach. Therefore, the results of this descriptive analysis led us to evaluate the prediction properties of the structural Theta, previously overlooked in the M4-competition.

Forecasts produced by the structural Theta are compared against several competing models, including all benchmarks included in the original M4-Competition, together with Auto Theta [29], the Dynamic Optimised Theta Model [9] and the Random Walk with Drift plus AutoRegressive model (RWDAR) recently introduced by [27]. Our findings clearly demonstrate the structural Theta's remarkable predictive capabilities, especially for trended series, as it consistently outperforms all its variants and competitors across the 100,000 series in M4. As such, the structural Theta provides a parsimonious yet strong benchmark to employ in business and economic forecasting. The remarkable performance of the structural Theta is further highlighted by the fact that the model can be easily estimated with minimal computational resources and time.

One might be concerned that the good performance of the structural Theta is specific to the forecasting application of the M4-Competition. However, the forecasting performance of the structural Theta model remains remarkably strong even when tested on the M5-competition dataset, which comprises hierarchical sales data, often characterized by intermittent and erratic patterns. This result is noteworthy given the significant difficulty in predicting such series using conventional forecasting models, such as those used in previous M competitions [21].

The rest of the paper proceeds as follows. Section 2 presents the structural Theta model and its estimation procedure. Section 3 focuses on some time-series properties (presence of drift and persistence) of the M4-dataset. Section 4 illustrates a Monte Carlo experiment designed for comparing the forecast accuracy of the structural Theta model relative to other benchmark models on simulated data, controlling for the presence of drift and persistence components. Section 5 evaluates the forecast accuracy of the structural Theta model on the M4-Competition dataset. Section 6 discusses a potential implementation of the structural Theta on the M5-Competition dataset. Lastly, Section 7 provides some final comments.

2. Model and estimation approach

2.1. The structural Theta model

The state-space representation of the structural Theta model is as follows:

$$\begin{aligned} y_t &= \mu_{t-1} + \epsilon_t \qquad \epsilon_t \sim \mathcal{NID}(0, \sigma^2_\epsilon) \\ \mu_t &= \omega + \mu_{t-1} + \eta_t \quad \eta_t \sim \mathcal{NID}(0, \sigma^2_\eta) \end{aligned} \tag{1}$$

where y_t is the observed time series, μ_t is its level (or stochastic trend), ω is a constant term, while $\mathcal{NID}$ denotes normally and independently distributed innovations and $t = 1, 2, \dots, n$ is the number of observations. It is also assumed that $E(\epsilon_t \eta_t) = 0$. In contrast to the single source of error state-space approach, the structural Theta model in (1) is driven by two uncorrelated noises, resulting in a formulation known as a MSOE, and belongs to the class of ”unobserved components time series models”, or ”structural time series models”, pioneered by [11].

Model (1) has an ARIMA(0,1,1) with drift reduced form:

$$(1 - L)y_t = \omega + (1 - L)\epsilon_t + \eta_{t-1} = \omega + v_t + \theta v_{t-1} \tag{2}$$

where L is the lag operator, such that $Ly_t = y_{t-1}$. The second equality in (2) holds true as far as the two expressions are observationally equivalent, as they share the same autocorrelation structure (i.e., an autocorrelation function which is zero after the first lag). Indeed the sum $(1 - L)\epsilon_t + \eta_{t-1}$ is a first-order moving average process, which can be rewritten as $v_t + \theta v_{t-1}$, where θ is the moving average coefficient and v_t is the innovation term. The corresponding forecast function is equal to a simple random walk update plus a constant linear trend: $y_{t+h} = \omega h + y_t \ \ (\forall h > 0)$, see [14], p. 289. In the empirical context, [22] and [26] provide substantial evidence supporting the applicability of the ARIMA(0,1,1) model to a variety of macroeconomic time series. Additionally, [33] demonstrates that temporal aggregation of a high-frequency random walk leads to an ARIMA(0,1,1) model with a positive-valued moving average coefficient.

Once fed into its state-space representation, model (1) can be estimated by the Kalman filter. The structural Theta features three parameters, however, as shown by [11], the model can be concentrated out by considering the "signal-to-noise" ratio: $q_\eta = \frac{\sigma^2_\eta}{\sigma^2_\epsilon}$. This reduces the number of parameters to be estimated and simplifies the Kalman filter recursions as follows (see [11], p. 107, Example 3.2.1):

- $\textbf{Innovation}: \quad v_t = y_t - m_{t-1}$
- $\textbf{Innovation variance}: \quad f_t = p_{t-1} + 1$
- $\textbf{Kalman gain}: \quad k_t = \frac{p_{t-1}}{p_{t-1}+1}$
- $\textbf{Prediction mean}: \quad m_t = E(\mu_t|Y_{t-1}) = \omega + m_{t-1} + k_t v_t$
- $\textbf{Prediction variance}: \quad p_t = E(\mu_t - m_t)^2 = p_{t-1} - \frac{p^2_{t-1}}{p_{t-1}+1} + q_\eta$

where $Y_{t-1} = \{y_1, y_2, \ldots, y_{t-1}\}$ denotes the information provided by data up to time $t-1$ and $Var(v_t|Y_{t-1}) = \sigma^2_\epsilon f_t$.

These expressions can be used to derive the prediction intervals (PI) for y_t at time $n+h$, $h = 1, 2, \ldots$. Those are:

$$PI(\hat{y}_{n+h|n}) = \hat{y}_{n+h|n} \pm z_{(1-\alpha)} \sqrt{MSE(\hat{y}_{n+h|n})} \tag{3}$$

where $\hat{y}_{n+h|n}$ is the h-step-ahead prediction of y_{n+h}, $z_{(1-\alpha)}$ is the quantile of a standard normal distribution at the desired significance level α and the forecast Mean Squared Error (MSE) is

$$MSE(\hat{y}_{n+h|n}) = \sigma^2_\epsilon (p_n + 1 + (h-1) q_\eta)$$

see [11], p. 148.

The model in (1) can be easily estimated by selecting the two parameters (q_η and ω) that maximize the *concentrated* log-likelihood function: $\log L_c = -\frac{n}{2} \log(2\pi + 1) -$

$\frac{1}{2}\sum_{t=1}^{n}\log(f_t) - \frac{n}{2}\log(\frac{1}{n}\sum_{t=1}^{n}\frac{v_t^2}{f_t})$ (see [11], p. 127, equation (3.4.9a)). Equivalently, one can select the two parameters that minimize $\sum_{t=1}^{n}\log(f_t) + n\log(\frac{1}{n}\sum_{t=1}^{n}\frac{v_t^2}{f_t})$.

The estimation comes at very low computational costs and no convergence issues. This aspect is a significant benefit of the structural Theta model, particularly when compared to more sophisticated and computationally intensive methods.

2.2. Seasonality in the structural Theta model

The model in (1) does not incorporate seasonality explicitly and consequently depends on an "external" seasonal adjustment procedure, in which the data are seasonally adjusted before conducting estimation and forecasting. Therefore, when deemed necessary based on the results of a standard seasonality test, we employ the conventional multiplicative decomposition approach. The same approach is used for the classic Theta method in the M4-Competition, see Table 2 in [20].

That is, defining s the seasons ($s \neq 1$), and assuming s is even, the smoothed series is given by $y_t^* = \sum_{j=-\frac{s}{2}}^{\frac{s}{2}} w_j y_{t-j}$ with $t = \frac{s}{2}+1,\ldots,n-\frac{s}{2}-1$ where $w_j = \frac{1}{s}$ for $j = 0, \pm1, ..., \pm(\frac{s}{2}-1)$ and $w_{-j} = w_j = \frac{1}{2s}$ for $j = \pm\frac{s}{2}$. The ratios $\frac{y_t}{y_t^*}$ are then averaged over each season to give a set of seasonal factors. Dividing the original series by these seasonal factors results in the seasonally adjusted series. The structural Theta forecasts carried out on the deseasonalised series are finally multiplied by the seasonal factors to obtain the seasonal forecasts.

3. The M4-Competition dataset

3.1. Overview

Forecast competitions have greatly influenced the field of forecasting over the years, providing a solid basis for assessing different approaches and learning empirically how to advance forecasting theory and practice [20].

The accuracy of the structural Theta model is tested on the M4-Competition dataset, which provides a comprehensive collection of time-series data from various domains, including finance, economics, demographics, and industry. In addition, the dataset covers a broad range of frequencies, such as yearly, quarterly, monthly, weekly, daily, along with hourly data, allowing for the assessment of forecasting methods across different time horizons.

Table 1 provides several insights into the M4 dataset, offering details on various aspects such as the distribution of the series across the different frequencies along with the forecasting horizons per data frequency. The yearly, monthly, and quarterly series collectively make up the vast majority of the dataset (95,000 time series), indicating their substantial influence on the overall outcomes of the forecast competition.

Each time series in the dataset is partitioned into a training set, used to estimate the different competing models, and a test/validation set that is employed to calculate forecast errors and assess the accuracy of the models. The number of forecasts required is 6 for yearly data, 8 for quarterly, 18 for monthly, 13 for weekly, 14 for daily and 48 for hourly.[5]

Table 1: Data used for the empirical evaluation (M4 dataset)

Frequency	Number of series	Forecasting horizon
Yearly	23,000	6
Quarterly	24,000	8
Monthly	48,000	18
Weekly	359	13
Daily	4,227	14
Hourly	414	48

[5]These forecasting horizons were chosen by the M4 organizers.

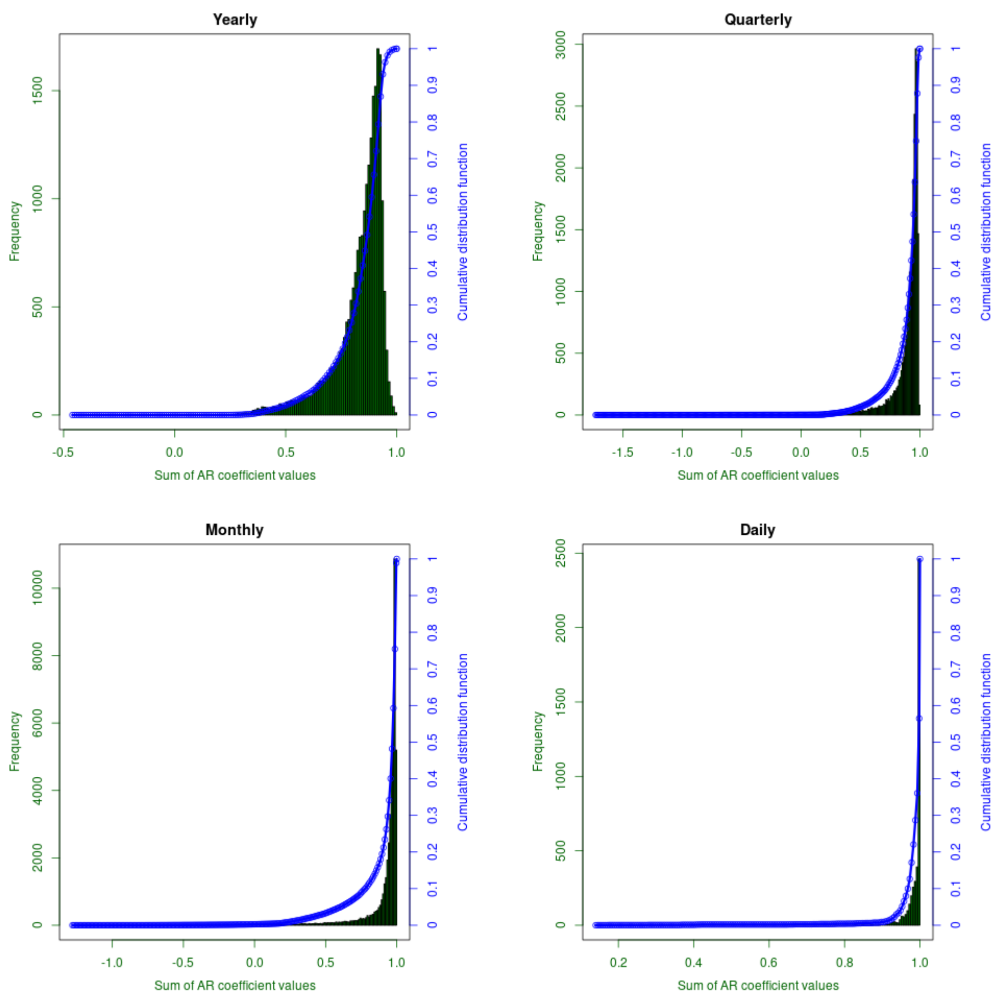
3.2. Persistence and trend significance in M4 series

The *ex-ante* evaluation of time-series properties is crucial for formulating a reliable forecasting model. Indeed, it ensures that the modelling approach aligns with the data characteristics. In this Section, we focus primarily on two time-series properties of the M4 dataset, namely, persistence and trend significance, see [30]. We claim that understanding these properties can help interpret the results of the forecast competition, also in light of the fact that the structural Theta model is well-suited for handling unit root data-generating processes. Although understanding the main time-series characteristics of such a large dataset as M4 is challenging, incorporating them into the forecasting framework is the key for successful forecasting.

The autoregressive (AR) process provides a clear, intuitive measure of persistence [30]. [1] suggest using the sum of the AR coefficients in order to assess the degree of persistence of a time series, based on the monotonic relation between the sum of the AR coefficients and the cumulative impulse response, which is a useful summary of the information contained in the impulse response function. Figure 1 thus illustrates the empirical distribution of the sum of the coefficients of an AR(p) process estimated on all the M4-Competition series in levels,[6] previously seasonally adjusted if deemed necessary based on the output of a standard seasonality test.

[6]The maximum order of the AR(p) process is set equal to the frequency of the data; for example, the maximum p is equal to 4 for quarterly data and to 12 for monthly data. The analysis has also been performed assuming an AR(1) process instead of an AR(p), resulting in very similar charts across all time frequencies, which can be provided by the authors upon request.

Figure 1: Distribution of the sum of the AR(p) coefficients estimated on the M4-Competition series across yearly, quarterly, monthly and daily frequencies



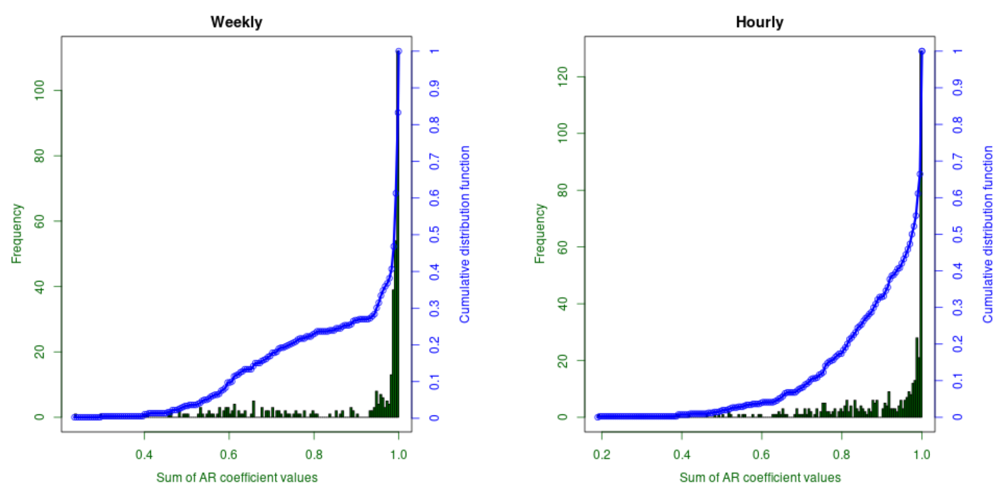
Notes. This figure shows the histograms and the empirical distribution functions of the sum of the AR(p) coefficients estimated on the time series of the M4-Competition across different frequencies (yearly, quarterly, monthly and daily).

Each chart refers to a single frequency (yearly, quarterly, monthly and daily) and shows the histogram with density values on left y-axis. The corresponding empirical cumulative distribution function is also added on the right y-axis in order to provide an alternative visualization of the sample distribution.

Focusing on yearly data, it can be seen that the empirical distribution of the sum of the AR coefficients is left-skewed, with the majority of values falling within the range [0.8, 1]. The same observation holds for the quarterly, monthly and daily frequencies, which tend to have an even higher concentration of values in this interval.

Even at weekly and hourly frequencies the distribution of the sum of the autoregressive coefficients is concentrated around one (Figure 2); yet, a non-negligible part of the probability distribution lies in the [0, 0.8] interval.

Figure 2: Distribution of the sum of the AR(p) coefficients estimated on the M4-Competition series across weekly and hourly frequencies



Notes. This figure shows the histograms and the empirical distribution functions of the sum of the AR(p) coefficients estimated on the time series of the M4-Competition across the remaining frequencies (weekly and hourly).

Overall, as also highlighted by [4], it is clear that the estimates of the sum of the AR(p) coefficients cluster near unity most of the times, pointing to high persistence in the data or near-unit root behavior.

A similar and stronger message arises when conducting unit root tests on the M4-

Competition time series. Specifically, Table 2 presents the results of augmented Dickey-Fuller tests, see [7], performed across all series for each time frequency, using a standard significance level of 5%. The null hypothesis of the test assesses whether the time series has a unit root as opposed to a stationary alternative hypothesis. The test can be executed in three versions: the first version with no drift and no linear trend, the second with drift but no linear trend, and the third with both drift and a linear trend. The results in Table 2 pertain to the second version.[7]

The findings reveal that a significant portion of the series exhibit a unit root, particularly at yearly, quarterly, monthly, and daily frequencies. The percentage of first-order integrated time series is also notably high at the weekly frequency. However, for hourly series, this proportion is closer to one-third. Overall, it is observed that more than 80% of the 100,000 time series in the M4-Competition dataset contain a unit root at the 5% significance level.

Table 2: Percentage of series with a unit root across the M4-Competition series

Frequency	% series with a unit root
Yearly	94.857
Quarterly	88.892
Monthly	76.467
Daily	95.481
Weekly	64.067
Hourly	32.126

Notes. This table reports the percentage of series for which the null hypothesis of a unit root is not rejected at the 5% significance level, based on the output of an augmented Dickey-Fuller test; see [7] and [6].

A closely related issue concerns the presence of a drift in the time series. Again, we

[7]Some preliminary ADF regressions have been tried on a number of selected time series with an intercept and a linear time trend. Most of the times the linear trend is not found to be significant at the 5% statistical level.

consider the standard augmented Dickey-Fuller equation with an intercept added to the test regression. Then, for those series for which the null hypothesis of a unit root is not rejected, it is possible to test the significance of the drift term given that the unit root is present, see [8] p. 223 (Table 4.1). The outcome of these t-tests ($\tau_{\alpha\mu}$) is summarized in Table 3 here below.

Table 3: Percentage of series with a drift given a unit root across the M4-Competition series

Frequency	% series with a drift given a unit root
Yearly	51.581
Quarterly	45.631
Monthly	33.801
Daily	19.530
Weekly	36.271
Hourly	21.452

Notes. This table reports the percentage of series for which the null hypothesis of no drift given a unit root is rejected at the 5% significance level, based on the output of an augmented Dickey-Fuller test; see [7] and [6].

The percentage of series with a drift, given a unit root, is quite high at the yearly and quarterly frequencies, where about half of those with a unit root also have a drift. The percentage drops to one-third for the monthly series. Overall, however, it emerges that more than 40% of the integrated series also have a drift at the 5% significance level. This share is anything but marginal. This evidence represents the starting point that led us to further investigate the prediction performance of the structural Theta model in the framework of the M4-Competition.

4. Monte Carlo study

This Section presents a Monte Carlo simulation that evaluates the forecasting performance of the structural Theta model while controlling for the presence of the deterministic drift and persistence in the data generating process (henceforth DGP).[8]

In this Monte Carlo experiment, data are simulated assuming either a trend stationary process (DGP 1) or a difference stationary process (DGP 2), following the seminal work of [22]. These two classes of non-stationary processes imply different dynamics and forecasts, thereby allowing us to validate the hypothesis of better performance of the structural Theta in the presence of persistence and deterministic drift components in the time series:

$$
\begin{array}{lll}
DGP\ 1 & DGP\ 2 & \\
y_t = \alpha + \beta t + c_t & (1-L)y_t = \beta + d_t & t = 1, 2, \ldots, n \\
(1-\phi L)c_t = (1+\theta L)u_t & (1-\delta L)d_t = (1+\lambda L)u_t & u_t \sim \mathcal{NID}(\mu_u, \sigma_u^2)
\end{array}
\quad (4)
$$

In DGP 1, α and β are fixed parameters, while c_t follows a stationary ARMA(1,1) process. DGP 1 exhibits a deterministic trend over time, but the fluctuations around this trend are stationary; the trend component can be removed through differencing, resulting in a stationary residual series. The variance of the series remains constant over time once the trend is removed.

In DGP 2, y_t is taken in first differences and d_t is a stationary ARMA(1,1). Therefore, DGP 2 accumulates an ARMA(1,1) over time, resulting in a degree of persistence

[8] The code developed for this simulation is available from the authors upon request.

clearly higher than that of DGP 1.[9] Differencing y_t removes both the trend and any other deterministic components, rendering the series stationary. Unlike DGP 1, DGP 2 may exhibit time-varying variance.

The sample size is 60 observations and the number of replications is 10,000.[10] All noises follow a Gaussian distribution. Data are simulated assuming random parameters. That is, for each replication, α, β and all the autoregressive and moving average coefficients are independently drawn from Uniform distributions. The use of random – rather than fixed – parameters allows the simulation experiment to be general rather than case-specific.

For each simulated series, six out-of-sample values are used as hold-out period to compare the forecast performance of the structural Theta model to ETS and auto ARIMA [13], two standard statistical benchmark methods also used in the original M4-Competition. Both point forecasts and prediction intervals are considered.

Results are presented in Table 4, with the top part showing point forecasts and the bottom part showing prediction intervals. Point forecast are compared using the Root Mean Squared Scaled Error (RMSSE), see [15]. Prediction intervals are evaluated using the Mean Scaled Interval Score (MSIS) metrics introduced by [10].

[9]DGP 2 can also be casted in state-space form:

$$\begin{aligned} y_t &= \mu_t + \chi_t \\ \mu_t &= \mu_{t-1} + \beta \\ \chi_t &= \chi_{t-1} + d_t \end{aligned} \tag{5}$$

and d_t follows a stationary ARMA(1,1) process.

[10]The overall results were not significantly affected by variations in the sample size.

Table 4: Monte Carlo averages of RMSSEs for point forecasts and MSISs for prediction intervals constructed using the competing methods

	DGP 1			DGP 2		
	Root Mean Squared Scaled Error (RMSSE)					
	MSOE Theta	ETS	AutoARIMA	MSOE Theta	ETS	AutoARIMA
h = 1	**0.716**	0.728	0.729	0.778	0.776	**0.771**
h = 2	**0.840**	0.856	0.857	**1.033**	1.040	1.038
h = 3	**0.914**	0.935	0.933	**1.234**	1.253	1.253
h = 4	**0.972**	0.999	0.994	**1.404**	1.438	1.437
h = 5	**1.020**	1.054	1.046	**1.563**	1.613	1.612
h = 6	**1.062**	1.103	1.090	**1.709**	1.779	1.776
	Mean Scaled Interval Score (MSIS)					
	MSOE Theta	ETS	AutoARIMA	MSOE Theta	ETS	AutoARIMA
85%	**5.554**	5.957	5.729	**9.465**	9.983	9.661
95%	**7.115**	7.569	7.294	12.686	12.595	**12.593**

Notes. The top part of this table refers to point forecasts, while the bottom part to prediction intervals. The benchmark methods used for comparison purposes are ETS and auto ARIMA. The most accurate method in terms of RMSSE/MSIS is highlighted in boldface.

In terms of point forecasts, the structural Theta model tends to outperform the other two competitors across all forecast horizons (except for $h = 1$ with DGP 2, when the three models have an almost equal performance). Regarding prediction intervals, the structural Theta is more accurate than both ETS and auto ARIMA, except for DGP 2 at 95%. These findings confirm that the structural Theta generally outperforms standard and flexible benchmarks such as ETS and auto ARIMA whenever data are generated by random walks or by stationary processes fluctuating around a deterministic trend.

5. Empirical evaluation using the M4-Competition data

The previous Section has shown that the structural Theta model outperforms standard statistical benchmarks when using simulated data with varying degrees of persistence and drift components. We now turn to the analysis of real-world time series and compare the performance of the structural Theta with the main benchmark models used in the M4-Competition dataset. These benchmarks encompass SES, Holt, Damped, and the Comb model, which corresponds to the simple arithmetic average of the first three models and was used for comparing all of the submitted methods; see Table 2 on page 57 in [20]. Furthermore, besides the classic Theta method, several alternative versions and extensions of it are considered, like Auto Theta [29], the Dynamic Optimized Theta Model [9], and the Random Walk with Drift plus AutoRegressive model (RWDAR), see [27].

The focus is, initially, on point forecasts. We assess the prediction performance using the same metrics adopted for the M4-competition (see [20]), that is, the symmetric mean absolute percentage error (or sMAPE, introduced by [17]), the mean absolute scaled error (or MASE, proposed by [15]) and the Overall Weighted Average (OWA), which corresponds to the average of the previous two.

5.1. Results for the yearly, quarterly, monthly and daily series

Results for the yearly, quarterly, monthly and daily series are presented in Table 5. At seasonal frequencies, whenever deemed necessary based on the output of a standard seasonality test, models are estimated on the seasonally adjusted data and then the final forecasts are reseasonalised following the procedure outlined in Section 2.2. Each column refers to a model, while each row reports the sMAPE, MASE, OWA values for evaluating the model's forecast accuracy. The overall ranking obtained by applying the OWA metric (calculated by expressing a model's sMAPE and MASE as a ratio of sMAPE and MASE achieved by the naive/random walk forecast) is also presented. The best result in terms of OWA is marked in bold.

Focusing on the yearly data, the structural Theta stands up as the most accurate model, followed in the second position by the RWDAR [27], which is akin to a random walk with drift plus an autoregressive term, and therefore shares many similarities with model (1). Unreported results show that the structural Theta model exhibits very similar predictive performance to the SSOE Theta variant, as largely anticipated, given their observational equivalence. This finding holds true in general also for the other time frequencies, not only for annual data.

The two extensions of the classic Theta (DOTM and Auto Theta) are positioned in the third and fourth ranking. The DOTM performs very similarly to RWDAR, while Auto Theta lags far behind. The Comb model ranks fifth, and the classic Theta method follows immediately after, with an OWA equal to 0.872 (9% higher compared to the structural Theta). All the other models/methods (SES, Holt and Damped) exhibit significantly lower performance.

The quarterly dataset conveys a similar message. The structural Theta model yields the lowest OWA, followed by DOTM and Auto Theta at a greater distance, and then by the Comb method, which holds the third position. Both for yearly and quarterly series, the structural Theta seems clearly superior in terms of forecast accuracy compared to the Theta classic.

Turning to monthly data, the ranking is slightly different since the structural Theta model comes in third position (0.916), anticipated by Auto Theta and the Theta classic. These results align closely with those presented in Tables 2 and 3. Indeed, while for yearly and quarterly frequencies we observe a strong presence of both persistence and drift (nearly one series out of two), for monthly series we detect the same in only about one series out of three. However, despite not being the top-performing model, the structural Theta model still displays strong forecasting capabilities when applied to monthly time series.

Table 5: Forecasting performance in terms of point forecasts across the Yearly, Quarterly, Monthly and Daily M4 series

Methods	MSOE Theta	SES	Holt	Damped	Comb	Theta classic	DOTM	Auto Theta	RWDAR
					Yearly				
sMAPE	**13.507**	16.398	16.535	15.162	14.874	14.603	13.677	13.797	13.690
MASE	**3.054**	3.981	3.576	3.372	3.282	3.382	3.075	3.184	3.059
OWA	**0.798**	1.003	0.956	0.888	0.868	0.872	0.805	0.823	0.804
Rank (OWA)	**1**	9	8	7	5	6	3	4	2
					Quarterly				
sMAPE	**10.023**	10.600	10.955	10.243	10.197	10.312	10.089	10.125	10.312
MASE	**1.165**	1.340	1.199	1.175	1.174	1.232	1.185	1.179	1.181
OWA	**0.880**	0.970	0.935	0.893	0.891	0.917	0.890	0.890	0.899
Rank (OWA)	**1**	8	7	4	3	6	2	2	5
					Monthly				
sMAPE	13.271	13.618	14.828	13.473	13.434	13.003	13.321	**13.125**	13.368
MASE	0.969	1.020	1.010	0.972	0.966	0.970	0.977	**0.960**	0.969
OWA	0.916	0.951	0.989	0.924	0.920	0.907	0.921	**0.906**	0.919
Rank (OWA)	3	8	9	7	5	2	6	**1**	4
					Daily				
sMAPE	3.014	3.045	3.070	3.063	**2.985**	3.053	3.042	3.025	3.064
MASE	3.169	3.226	3.169	3.179	**3.148**	3.207	3.209	3.176	3.215
OWA	0.987	1.000	0.996	0.996	**0.978**	0.999	0.997	0.989	1.002
Rank (OWA)	2	7	4	4	**1**	6	5	3	8

Notes. The results are presented per data frequency. The most accurate method in terms of OWA is highlighted in boldface.

Moreover, our proposed approach ranks second using daily series, again, outperforming all variants of Theta.

5.2. *Weekly and hourly time series*

On the other hand, for both hourly and weekly series it is evident that the performance of the structural Theta model deteriorates, as shown in Table 6. This finding is in line with the properties described in Section 3.2, with particular reference to the integration order and presence of trends.

Table 6: Forecasting performance in terms of point forecasts across the Weekly and Hourly M4 series

Methods	MSOE Theta	SES	Holt	Damped	Comb	Theta classic	DOTM	Auto Theta	RWDAR
					Weekly				
sMAPE	8.939	9.012	9.706	**8.867**	8.947	9.094	9.089	8.987	8.548
MASE	2.547	2.685	2.413	**2.400**	2.429	2.637	2.583	2.488	2.543
OWA	0.946	0.975	0.964	**0.916**	0.926	0.971	0.961	0.938	0.924
Rank (OWA)	5	9	7	**1**	3	8	6	4	2
					Hourly				
sMAPE	18.661	18.094	29.474	19.277	22.114	18.138	18.093	17.754	**14.093**
MASE	3.418	2.385	9.380	2.947	4.585	2.455	2.388	1.807	**2.253**
OWA	1.221	0.990	2.760	1.140	1.559	1.006	0.991	0.860	**0.854**
Rank (OWA)	7	3	9	6	8	5	4	2	**1**

Notes. The results are presented per data frequency. The most accurate method in terms of OWA is highlighted in boldface.

5.3. *Overall performance*

We now evaluate the average performance of the models across all the 100,000 M4 time series. Table 7 clearly shows that the structural Theta model outperforms all competitors considered in our forecasting exercise. Additionally, the structural Theta model performs very well even among all the 61 methods considered for point forecasts in the

M4-Competition.[11] Indeed, in terms of OWA, it ranks in the 10th position (we refer to Table 4 in [20]). Noteworthy, the structural Theta model outperforms all benchmarks (classic Theta, Comb, Damped, Holt, SES, Nave 2, Nave 1, Nave S, RNN, MLP) and two standard methods (ARIMA, ETS). It also beats sophisticated extensions of the classic Theta method, such as Auto Theta and DOTM. We believe that these points emphasize the significance of our contribution.

How can such a simple and parsimonious model consistently outperform many statistically sophisticated competitors, even ranking among the top performers? We believe that the answer lies in the descriptive analysis carried out in Section 3.2, which lays the groundwork for the subsequent forecast evaluation. As observed, the prevalence of unit roots and drifts in the series strongly supports our simple forecasting approach. Indeed, these two factors seem to be particularly relevant for successful out-of-sample forecasting.

Table 7: Forecasting performance in terms of point forecasts across all the M4-Competition series

Methods	MSOE Theta	SES	Holt	Damped	Comb	Theta classic	DOTM	Auto Theta	RWDAR
	All frequencies								
sMAPE	**12.119**	13.088	13.836	12.654	12.566	12.312	12.197	12.137	12.259
MASE	**1.604**	1.883	1.776	1.679	1.661	1.694	1.615	1.627	1.606
OWA	**0.866**	0.975	0.975	0.906	0.898	0.897	0.872	0.873	0.872
Rank (OWA)	**1**	7	7	6	5	4	2	3	2

Notes. The most accurate method in terms of OWA is highlighted in boldface.

[11] These 61 competitors include 49 submitted methods, 10 benchmarks, and two standard methods for comparison.

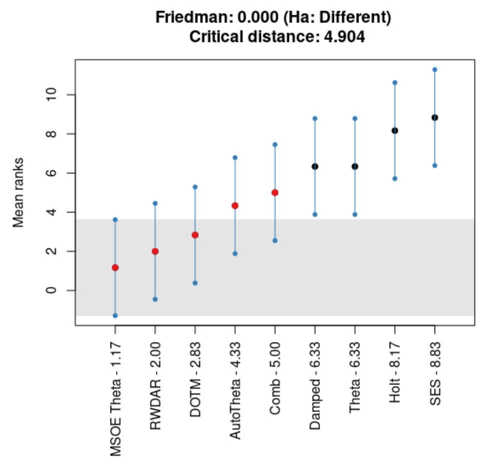
5.4. Multiple comparisons with the best (MCB)

In this Section, we implement the nonparametric Nemenyi test [23], which is comparable to the multiple-comparisons procedure described by [16] (multiple comparisons with the best method, MCB). This test assesses the statistical significance of observed accuracy differences between forecasts; see [5] and [12]. It ranks the performance of methods for each time series, computes the means of those ranks (medians), and produces confidence intervals for those means. If the confidence intervals for different methods overlap, it indicates that the means are not significantly different. Otherwise, it identifies which method has a higher rank and which has a lower one.

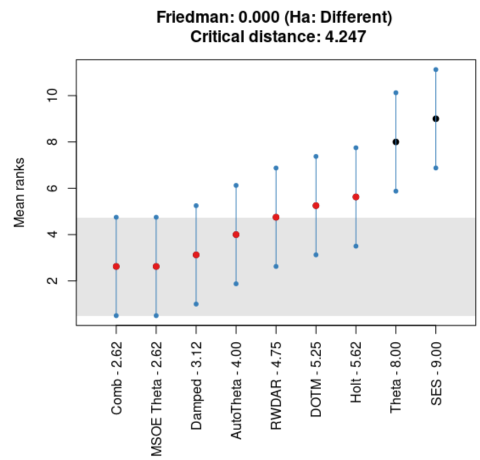
The Nemenyi test is conducted separately for yearly, quarterly, monthly, and daily series from the M4-Competition dataset, employing the Mean Absolute Scaled Error (MASE) metric as used by [29], and the `nemenyi` function in the `tsutils` package for R. The results of the Nemenyi test are presented in Figure 3 for each time frequency.

Figure 3 shows that at all frequencies the Friedman test rejects the null hypothesis of no difference in accuracy among the forecast models. Thus, post-hoc analysis based on the Nemenyi test is performed to evaluate which differences are significant. For yearly data, the structural Theta model exhibits the best-ranked performance, which is not statistically different from that of RWDAR, DOTM, Auto Theta, and Comb. However, the structural Theta model statistically outperforms all other methods above the reference value without any overlap with the shaded area. A somewhat similar ranking emerges for quarterly data, with the structural Theta model ranked first along with Comb. For monthly data, the structural Theta is no longer the best model, but its performance is not statistically different from the best forecasting method (Auto Theta). Even with daily data, the structural Theta ranks third and consistently outperforms classic Theta, DOTM, RWDAR, and SES.

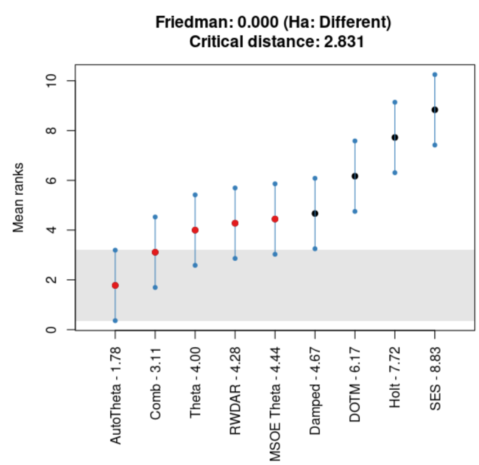
Figure 3: Comparison of forecast models using Multiple Comparisons with the Best Method



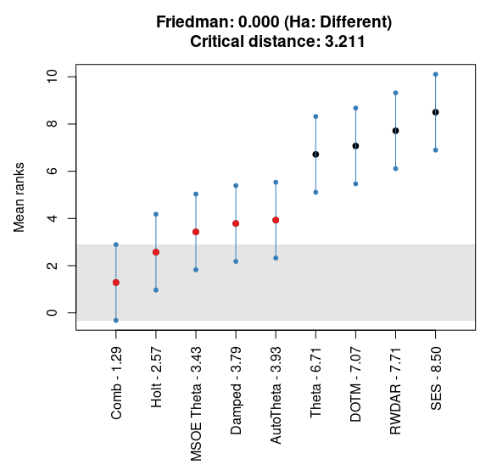
(a) Yearly series



(b) Quarterly series



(c) Monthly series



(d) Daily series

Notes. This figure shows the Nemenyi test using MASE as error measure separately for yearly, quarterly, monthly and daily series. The confidence level used for the comparison is 5%. The models are ordered based on their average ranks.

Drawing overall conclusions from this analysis, the previous results described in Section 5.1 are broadly confirmed: the structural Theta model is found to perform well compared to other benchmark models in the M4-Competition dataset, with significant improvements upon strong benchmark models and a stable performance across the main time frequencies.

5.5. Prediction intervals

So far the predictive properties of the structural Theta model have been evaluated in the M4 dataset focusing on point forecasts. This Section turns to prediction intervals and probabilistic performance metrics.

Table 8 examines the performance of the structural Theta model with respect to the average Mean Scaled Interval Score (MSIS) introduced by [10], which has also been employed in the M4-Competition for ranking interval forecasts. The benchmark methods used for comparison purposes are Naive, SES, Holt, Damped, DOTM and Auto Theta.

The results demonstrate that the structural Theta model significantly outperforms Naive, SES, and Holt benchmarks across yearly, quarterly, and monthly frequencies. However, at the yearly frequency, Auto Theta emerges as the most accurate method, outperforming the structural Theta. For quarterly data, Damped ranks first, followed closely by Auto Theta and the structural Theta. Conversely, the structural Theta model exhibits the highest accuracy at the monthly frequency, with DOTM and Auto Theta following closely behind.

Overall, we conclude that the structural Theta model produces accurate forecasts not only in terms of point forecasts but also in terms of prediction intervals, despite some signs of underestimation of uncertainty, particularly when forecasting yearly time series.

Table 8: Forecasting performance in terms of precision of prediction intervals across the Yearly, Quarterly and Monthly M4-Competition series

	Yearly	Quarterly	Monthly
Naive	56.554	14.073	12.300
SES	56.038	13.595	10.923
Holt	47.427	11.833	11.385
Damped	43.419	**11.258**	10.439
MSOE Theta	46.777	11.741	**9.576**
DOTM	48.847	12.255	9.835
Auto Theta	**33.127**	11.464	9.830

Notes. The performances of the prediction intervals are evaluated using the MSIS. For each frequency, the most accurate method in terms of MSIS is highlighted in boldface. A 95% prediction interval framework for estimating the uncertainty around the point forecasts has been adopted.

6. Empirical evaluation using the M5-Competition data

One could raise concerns about whether the excellent forecasting performance of the structural Theta model is solely attributable to its application in the M4-competition dataset. To address this issue, we decided to evaluate its prediction performance in the M5-competition, which followed the M4-competition and exclusively focused on retail sales forecasting. The M5 dataset contains 42,840 hierarchical sales records from Walmart, covering stores in three US states (California, Texas, and Wisconsin). It includes item-level, department, product category, and store details for a period of 5 years, starting from January 29, 2011, to April 24, 2016. The goal of the M5 competition is to forecast daily sales for the next 28 days, until May 22, 2016; see [21] for additional details. The hierarchical aggregation structure of the dataset makes it feasible to produce forecasts using a bottom-up approach or a top-down approach. Specifically, it is possible to forecast the series at the most disaggregated level and derive the aggregated forecasts using a bottom-up approach, or rather forecast the most aggregated series and compute the remaining series using proportions (top-down approach).

One of the main features of the M5-competition dataset is the widespread presence of zeros. This presents a challenge for standard time-series models, as they can hardly accommodate the presence of many zeros in their specifications. Despite this challenge, Table 9 reports the performance achieved by the structural Theta model in terms of point forecasts, as measured by both the overall accuracy and across the 12 aggregation levels, using the Weighted Root Mean Squared Scaled Error (WRMSSE) metrics. Table 9 also presents the performance of the structural Theta model in terms of prediction intervals, based on the Weighted Scaled Pinball Loss (WSPL) metrics.

Regarding point forecasts, the structural Theta model provides accurate predictions when compared to other standard benchmarks, such as Naive, seasonal Naive, simple exponential smoothing (SES), ARIMA, ARIMAX, etc.[12] On average, the top-down (TD) structural Theta ranks fifth among the 24 benchmark methods, while the bottom-up (BU) structural Theta ranks third, even outperforming ARIMAX models and combinations. This is not surprising given that the best performing benchmark (Exponential Smoothing bottom-up) is a similar method, except for the presence of a drift.

However, the most relevant results pertain to the uncertainty competition (WSPL measure). Our approach outperforms all six benchmarks, including Naive and seasonal Naive, ETS, SES, ARIMA and empirical quantile estimation (Kernel).

While these results are promising, they should be interpreted with caution as the structural Theta model is not primarily designed to handle intermittent data. Nevertheless, they demonstrate that the structural Theta model can perform reasonably well even when applied to retail sales from the M5-Competition dataset, which exhibit intermittency.

[12]We refer to the complete set of results and benchmarks used in the M5-Competition as reported at `https://github.com/Mcompetitions/M5-methods`

Table 9: Forecasting performance of the structural Theta in the M5-competition in terms of WRMSSE (accuracy) and WSPL (uncertainty)

	Aggregation levels												Average
	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	
	Accuracy competition (WRMMSE)												
TD	0.442	0.538	0.609	0.478	0.562	0.603	0.700	0.665	0.753	1.043	0.984	0.916	0.691
BU	0.443	0.531	0.597	0.489	0.564	0.590	0.662	0.665	0.736	1.016	0.971	0.916	0.681
	Uncertainty competition (WSPL)												
TD	0.144	0.149	0.168	0.151	0.173	0.166	0.192	0.181	0.207	0.314	0.297	0.302	0.203

Notes. The results are presented at both the aggregation level, as described in [21], and overall. "Average" is the average of all aggregation levels.

7. Conclusions

The Theta stands as a well-established forecasting method, having found widespread application in forecast competitions over the past two decades. This paper discusses the MSOE version of the Theta method in its state-space form, belonging to the family of structural time series models. The proposed model corresponds to a random walk plus constant with an additional error term and its forecast function is equal to a random walk update plus a constant linear trend. An extensive examination of its out-of-sample forecasting capabilities has been conducted using the M4-competition dataset.

Results highlight that the structural Theta model provides a simple yet effective approach to forecasting, especially when the data exhibit both persistence and drift components. Indeed, using the M4-dataset, this model consistently outperforms all the Theta variants and other competitors across 100,000 time series. These results are confirmed by simulations. Even when evaluated on the M5-competition dataset, which consists of hierarchical sales data known for its intermittent and unpredictable patterns, the structural Theta model continues to exhibit remarkably strong forecasting performance. Thus, given its notable forecasting accuracy and robust performance in M4 and M5, the structural Theta emerges as a highly suitable model and a reference point to serve as a benchmark in the context of forecast competitions. Another merit of the structural Theta model is that it is fully replicable and can be estimated with minimal computational effort. In light of this, we believe that this particular variant of the Theta model holds the potential for valuable application in economic and business forecasting.

References

[1] Donald W. K. Andrews and Hong-Yuan Chen. Approximately median-unbiased estimation of autoregressive models. *Journal of Business & Economic Statistics*, 12(2):187–204, 1994.

[2] Vassilis Assimakopoulos and Konstantinos Nikolopoulos. The theta model: a decomposition approach to forecasting. *International Journal of Forecasting*, 16(4): 521–530, 2000.

[3] Jos Casals, Sonia Sotoca, and Miguel Jerez. Single and multiple error state-space models for signal extraction. *Journal of Statistical Computation and Simulation*, 85(5):1053–1069, 2015.

[4] Jennifer L. Castle, Jurgen A. Doornik, and David F. Hendry. Forecasting principles from experience with forecasting competitions. *Forecasting*, 3(1):138–165, 2021.

[5] Janez Demšar. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7(1):1–30, 2006.

[6] David Dickey and Wayne A Fuller. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4):1057–72, 1981.

[7] David A Dickey and Wayne A Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431, 1979.

[8] W. Enders. *Applied Econometric Times Series*. Wiley Series in Probability and Statistics - Applied Probability and Statistics Section Series. Wiley, 1995.

[9] Jose A. Fiorucci, Tiago R. Pellegrini, Francisco Louzada, Fotios Petropoulos, and Anne B. Koehler. Models for optimising the theta method and their relationship to state space models. *International Journal of Forecasting*, 32(4):1151–1161, 2016.

[10] Tilmann Gneiting and Adrian E Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.

[11] Andrew C. Harvey. *Forecasting, Structural Time Series Analysis, and the Kalman Filter*. Cambridge, 1989.

[12] M. Hollander, D.A. Wolfe, and E. Chicken. *Nonparametric Statistical Methods*. Wiley Series in Probability and Statistics. Wiley, 2013.

[13] Rob Hyndman and Yeasmin Khandakar. Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software*, 27(03):1–22, 2008.

[14] Rob J. Hyndman and Baki Billah. Unmasking the Theta method. *International Journal of Forecasting*, 19(2):287–290, 2003.

[15] Rob J. Hyndman and Anne B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688, 2006.

[16] Alex J. Koning, Philip Hans Franses, Michele Hibon, and H.O. Stekler. The M3 competition: Statistical tests of the results. *International Journal of Forecasting*, 21(3):397–409, 2005.

[17] Spyros Makridakis. Accuracy measures: theoretical and practical concerns. *International Journal of Forecasting*, 9(4):527–529, December 1993.

[18] Spyros Makridakis and Michèle Hibon. The m3-competition: results, conclusions and implications. *International Journal of Forecasting*, 16(4):451–476, 2000. The M3-Competition.

[19] Spyros Makridakis, Evangelos Spiliotis, and Vassilios Assimakopoulos. Predicting/hypothesizing the findings of the m4 competition. *International Journal of Forecasting*, 36(1):29–36, 2020. M4 Competition.

[20] Spyros Makridakis, Evangelos Spiliotis, and Vassilios Assimakopoulos. The m4 competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*, 36(1):54–74, 2020.

[21] Spyros Makridakis, Evangelos Spiliotis, and Vassilios Assimakopoulos. The m5 competition: Background, organization, and implementation. *International Journal of Forecasting*, 38(4):1325–1336, 2022.

[22] Charles R. Nelson and Charles I. Plosser. Trends and random walks in macroeconmic time series: Some evidence and implications. *Journal of Monetary Economics*, 10(2):139–162, 1982.

[23] Peter Bjorn Nemenyi. *Distribution-free multiple comparisons.* Princeton University, 1963.

[24] Kostas I. Nikolopoulos and Dimitrios D. Thomakos. *Forecasting with the Theta Method: Theory and Applications*. John Wiley and Sons, Ltd, 2019.

[25] Kum Hwa Oh, Eric Zivot, and Drew Creal. The relationship between the Beveridge-Nelson decomposition and other permanent-transitory decompositions that are popular in economics. *Journal of Econometrics*, 146(2):207–219, October 2008.

[26] Robert Rossana and John Seater. Temporal aggregation and economic time series. *Journal of Business & Economic Statistics*, 13(4):441–51, 1995.

[27] Giacomo Sbrana and Andrea Silvestrini. The rwdar model: A novel state-space approach to forecasting. *International Journal of Forecasting*, 39(2):922–937, 2023. ISSN 0169-2070.

[28] Evangelos Spiliotis, Vassilios Assimakopoulos, and Konstantinos Nikolopoulos. Forecasting with a hybrid method utilizing data smoothing, a variation of the theta method and shrinkage of seasonal factors. *International Journal of Production*

Economics, 209:92–102, 2019. The Proceedings of the 19th International Symposium on Inventories.

[29] Evangelos Spiliotis, Vassilios Assimakopoulos, and Spyros Makridakis. Generalizing the theta method for automatic forecasting. *European Journal of Operational Research*, 284(2):550–558, 2020.

[30] Evangelos Spiliotis, Andreas Kouloumos, Vassilios Assimakopoulos, and Spyros Makridakis. Are forecasting competitions data representative of the reality? *International Journal of Forecasting*, 36(1):37–53, 2020. M4 Competition.

[31] Dimitrios D. Thomakos and Konstantinos Nikolopoulos. Forecasting multivariate time series with the theta method. *Journal of Forecasting*, 34(3):220–229, April 2015.

[32] Mark W. Watson. Univariate detrending methods with stochastic trends. *Journal of Monetary Economics*, 18(1):49–75, 1986.

[33] Holbrook Working. Note on the correlation of first differences of averages in a random chain. *Econometrica*, 28(4):916–918, 1960.

RECENTLY PUBLISHED "TEMI" (*)

N. 1431 – *Monetary and fiscal policy responses to fossil fuel price shocks*, by Anna Bartocci, Alessandro Cantelmo, Pietro Cova, Alessandro Notarpietro and Massimiliano Pisani (December 2023).

N. 1432 – *Do female leaders choose women? Evidence from visible and hidden appointments*, by Andrea Cintolesi and Edoardo Frattola (December 2023).

N. 1433 – *Monetary policy tightening in response to uncertain stagflationary shocks: a model-based analysis*, by Anna Bartocci, Alessandro Cantelmo, Alessandro Notarpietro and Massimiliano Pisani (December 2023).

N. 1434 – *Inflation, capital structure and firm value*, by Andrea Fabiani and Fabio Massimo Piersanti (December 2023).

N. 1435 – *Announcement and implementation effects of central bank asset purchases*, by Marco Bernardini and Antonio M. Conti (December 2023).

N. 1436 – *Connecting the dots: the network nature of shocks propagation in credit markets*, by Stefano Pietrosanti and Edoardo Rainone (December 2023).

N. 1437 – *Inflation expectations and misallocation of resources: evidence from Italy*, by Tiziano Ropele, Yuriy Gorodnichenko and Olivier Coibion (December 2023).

N. 1438 – *Women in economics: the role of gendered references at entry in the profession*, by Audinga Baltrunaite, Alessandra Casarico and Lucia Rizzica. (February 2024).

N. 1439 – *Procuring survival*, by Matilde Cappelletti, Leonardo M. Giuffrida and Gabriele Rovigatti. (February 2024).

N. 1440 – *Estimating the returns to occupational licensing: evidence from regression discontinuities at the bar exam*, by Omar Bamieh, Andrea Cintolesi and Mario Pagliero. (February 2024).

N. 1441 – *Household perceived sources of business cycle fluctuations: a tale of supply and demand*, by Clodomiro Ferreira and Stefano Pica. (February 2024).

N. 1442 – *Aggregate uncertainty, HANK, and the ZLB*, by Alessandro Lin and Marcel Peruffo (March 2024).

N. 1443 – *Monetary policy under natural disaster shocks*, by Alessandro Cantelmo, Nikos Fatouros, Giovanni Melina and Chris Papageorgiou (March 2024).

N. 1444 – *Endogenous job destruction risk and aggregate demand shortages*, by Nicolò Gnocato (March 2024).

N. 1445 – *Carbon taxes around the world: cooperation, strategic interactions, and spillovers*, by Alessandro Moro and Valerio Nispi Landi (March 2024).

N. 1446 – *Nowcasting Italian GDP growth: a Factor MIDAS approach*, by Donato Ceci, Orest Prifti and Andrea Silvestrini (March 2024).

N. 1447 – *The green sin: how exchange rate volatility and financial openness affect green premia*, by Alessandro Moro and Andrea Zaghini (March 2024).

N. 1448 – *Oil price shocks in real time*, by Andrea Gazzani, Fabrizio Venditti and Giovanni Veronese (March 2024).

N. 1449 – *Market perceptions, monetary policy, and credibility*, by Vincenzo Cuciniello (March 2024).

N. 1450 – *Energy price shocks, unemployment, and monetary policy*, by Nicolò Gnocato (March 2024).

N. 1451 – *The impact of hydrogeological events on firms: evidence from Italy*, Stefano Clò, Francesco David and Samuele Segoni (April 2024).

N. 1452 – *Measuring households' financial fragilities: an analysis at the intersection of income, financial wealth and debt*, by David Loschiavo, Federico Tullio and Antonietta di Salvatore (April 2024).

N. 1453 – *Unconventionally green*, Andrea Zaghini (April 2024).

(*) Requests for copies should be sent to:
Banca d'Italia – Servizio Studi di struttura economica e finanziaria – Divisione Biblioteca e Archivio storico – Via Nazionale, 91 – 00184 Rome – (fax 0039 06 47922059). They are available on the Internet www.bancaditalia.it.

2022

ANDINI M., M. BOLDRINI, E. CIANI, G. DE BLASIO, A. D'IGNAZIO and A. PALADINI, *Machine learning in the service of policy targeting: the case of public credit guarantees,* Journal of Economic Behavior & Organization, v. 198, pp. 434-475, **WP 1206 (February 2019).**

ANGELICO C., J. MARCUCCI, M. MICCOLI and F. QUARTA, *Can we measure inflation expectations using twitter?,* Journal of Econometrics, v. 228, 2, pp. 259-277, **WP 1318 (February 2021).**

BARTOCCI A., A. NOTARPIETRO and M. PISANI, *Covid-19 shock and fiscal-monetary policy mix in a monetary union,* Economic challenges for Europe after the pandemic, Springer Proceedings in Business and Economics, Berlin-Heidelberg, Springer, **WP 1313 (December 2020).**

BOTTERO M., C. MINOIU, J. PEYDRÒ, A. POLO, A. PRESBITERO and E. SETTE, *Expansionary yet different: credit supply and real effects of negative interest rate policy,* Journal of Financial Economics, v. 146, 2, pp. 754-778, **WP 1269 (March 2020).**

BRONZINI R., A. D'IGNAZIO and D. REVELLI, *Financial structure and bank relationships of Italian multinational firms,* Journal of Multinational Financial Management, v. 66, Article 100762, **WP 1326 (March 2021).**

CANTELMO A., *Rare disasters, the natural interest rate and monetary policy,* Oxford Bulletin of Economics and Statistics, v. 84, 3, pp. 473-496, **WP 1309 (December 2020).**

CARRIERO A., F. CORSELLO and M. MARCELLINO, *The global component of inflation volatility,* Journal of Applied Econometrics, v. 37, 4, pp. 700-721, **WP 1170 (May 2018).**

CIAPANNA E. and G. ROVIGATTI, *The grocery trolley race in times of Covid-19. Evidence from Italy,* Italian Economic Journal / Rivista italiana degli economisti, v. 8, 2, pp. 471-498, **WP 1341 (June 2021).**

CONTI A. M., A. NOBILI and F. M. SIGNORETTI, *Bank capital requirement shocks: a narrative perspective,* European Economic Review, v.151, Article 104254, **WP 1199 (November 2018).**

FAIELLA I. and A. MISTRETTA, *The net zero challenge for firms' competitiveness,* Environmental and Resource Economics, v. 83, pp. 85-113, **WP 1259 (February 2020).**

FERRIANI F. and G. VERONESE, *Hedging and investment trade-offs in the U.S. oil industry,* Energy Economics, v. 106, Article 105736, **WP 1211 (March 2019).**

GUISO L., A. POZZI, A. TSOY, L. GAMBACORTA and P. E. MISTRULLI, *The cost of steering in financial markets: evidence from the mortgage market,* Journal of Financial Economics, v.143, 3, pp. 1209-1226, **WP 1252 (December 2019).**

LAMORGESE A. and D. PELLEGRINO, *Loss aversion in housing appraisal: evidence from Italian homeowners,* Journal of Housing Economics, v. 56, Article 101826, **WP 1248 (November 2019).**

LI F., T. MÄKINEN, A. MERCATANTI and A. SILVESTRINI, *Causal analysis of central bank holdings of corporate bonds under interference,* Economic Modelling, v.113, Article 105873, **WP 1300 (November 2020).**

LOBERTO M, A. LUCIANI and M. PANGALLO, *What do online listings tell us about the housing market?,* International Journal of Central Banking, v. 18, 4, pp. 325-377, **WP 1171 (April 2018).**

MIRENDA L., M. SAURO and L. RIZZICA, *The economic effects of mafia: firm level evidence,* American Economic Review, vol. 112, 8, pp. 2748-2773, **WP 1235 (October 2019).**

MOCETTI S., G. ROMA and E. RUBOLINO, *Knocking on parents' doors: regulation and intergenerational mobility,* Journal of Human Resources, v. 57, 2, pp. 525-554, **WP 1182 (July 2018).**

PERICOLI M. and M. TABOGA, *Nearly exact Bayesian estimation of non-linear no-arbitrage term-structure models,* Journal of Financial Econometrics, v. 20, 5, pp. 807-838, **WP 1189 (September 2018).**

ROSSI P. and D. SCALISE, *Financial development and growth in European regions,* Journal of Regional Science, v. 62, 2, pp. 389-411, **WP 1246 (November 2019).**

SCHIVARDI F., E. SETTE and G. TABELLINI, *Credit misallocation during the European financial crisis,* Economic Journal, v. 132, 641, pp. 391-423, **WP 1139 (September 2017).**

TABOGA M., *Cross-country differences in the size of venture capital financing rounds: a machine learning approach,* Empirical Economics, v. 62, 3, pp. 991-1012, **WP 1243 (November 2019).**

2023

APRIGLIANO V., S. EMILIOZZI, G. GUAITOLI, A. LUCIANI, J. MARCUCCI and L. MONTEFORTE, *The power of text-based indicators in forecasting Italian economic activity,* International Journal of Forecasting, v. 39, 2, pp. 791-808, **WP 1321 (March 2021).**

BARTOCCI A., A. NOTARPIETRO and M. PISANI, *Non-standard monetary policy measures in non-normal times,* International Finance, v. 26, 1, pp. 19-35, **WP 1251 (November 2019).**

CAPPELLETTI G. and P. E. MISTRULLI, *The role of credit lines and multiple lending in financial contagion and systemic events,* Journal of Financial Stability, v. 67, Article 101141, **WP 1123 (June 2017).**

CECI D. and A. SILVESTRINI, *Nowcasting the state of the Italian economy: the role of financial markets,* Journal of Forecasting, v. 42, 7, pp. 1569-1593, **WP 1362 (February 2022).**

CIAPANNA E, S. MOCETTI and A. NOTARPIETRO, *The macroeconomic effects of structural reforms: an empirical and model-based approach,* Economic Policy, v. 38, 114, pp. 243-285, **WP 1303 (November 2020).**

CORNELI F., *Sovereign debt maturity structure and its costs,* International Tax and Public Finance, v. 31, 1, pp. 262-297, **WP 1196 (November 2018).**

DAURICH D, S. DI ADDARIO and R. SAGGIO, *The macroeconomic effects of structural reforms: an empirical and model-based approach,* Review of Economic Studies, v. 90, 6, pp. 2880–2942, **WP 1390 (November 2022).**

DI ADDARIO S., P. KLINE, R. SAGGIO and M. SØLVSTEN, *The effects of partial employment protection reforms: evidence from Italy,* Journal of Econometrics,v. 233, 2, pp. 340-374, **WP 1374 (June 2022).**

FERRARI A. and V. NISPI LANDI, *Toward a green economy: the role of central bank's asset purchases,* International Journal of Central Banking, v. 19, 5, pp. 287-340, **WP 1358 (February 2022).**

FERRIANI F., *Issuing bonds during the Covid-19 pandemic: was there an ESG premium?,* International Review of Financial Analysis, v. 88, Article 102653, **WP 1392 (November 2022).**

GIORDANO C., *Revisiting the real exchange rate misalignment-economic growth nexus via the across-sector misallocation channel,* Review of International Economics, v. 31, 4, pp. 1329-1384, **WP 1385 (October 2022).**

GUGLIELMINETTI E., M. LOBERTO and A. MISTRETTA, *The impact of COVID-19 on the European short-term rental market,* Empirica, v. 50, 3, pp. 585-623, **WP 1379 (July 2022).**

LILLA F., *Volatility bursts: a discrete-time option model with multiple volatility components,* Journal of Financial Econometrics, v. 21, 3, pp. 678-713, **WP 1336 (June 2021).**

LOBERTO M., *Foreclosures and house prices,* Italian Economic Journal / Rivista italiana degli economisti, v. 9, 1, pp. 397-424, **WP 1325 (March 2021).**

LOMBARDI M. J., M. RIGGI and E. VIVIANO, *Worker's bargaining power and the Phillips curve: a micro-macro analysis, and wages,* Journal of the European Economic Association, v. 21, 5, pp. 1905–1943, **WP 1302 (November 2020).**

NERI S., *Long-term inflation expectations and monetary policy in the Euro Area before the pandemic*, European Economic Review, v. 154, Article 104426, **WP 1357 (December 2021).**

ORAME A., *Bank lending and the European debt crisis: evidence from a new survey*, International Journal of Central Banking, v. 19, 1, pp. 243-300, **WP 1279 (June 2020).**

RIZZICA L., G. ROMA and G. ROVIGATTI, *The effects of shop opening hours deregulation: evidence from Italy,* The Journal of Law and Economics, v. 66, 1, pp. 21-52, **WP 1281 (June 2020).**

TANZI G. M., *Scars of youth non-employment and labour market conditions,* Italian Economic Journal / Rivista italiana degli economisti, v. 9, 2, pp. 475-499, **WP 1312 (December 2020).**

2024

BRANZOLI N., R. GALLO, A. ILARI and D. PORTIOLI, *Central banks' corporate asset purchase programmes and risk-taking by bond funds in the aftermath of market stress,* Journal of Financial Stability, v. 72, Article 101261, **WP 1404 (March 2023).**

BRANZOLI N., E. RAINONE and I. SUPINO, *The role of banks' technology adoption in credit markets during the pandemic,* Journal of Financial Stability, v. 71, Article 101230, **WP 1406 (March 2023).**

BUONO I, F. CORNELI and E. DI STEFANO, *Capital inflows to emerging countries and their sensitivity to the global financial cycle,* International Finance, v. 27,2, pp. 17-34, **WP 1262 (February 2020).**

DEL PRETE S, G. PAPINI and M. TONELLO, *Gender quotas, board diversity and spillover effects. Evidence from Italian banks,* Journal of Economic Behavior & Organization, v. 221, pp. 148-173, **WP 1395 (December 2022).**

DE MARCHI R. and A. MORO, *Forecasting fiscal crises in emerging markets and low-income countries with machine learning models,* Open Economies Review. v. 35, 1, pp. 189-213, **WP 1405 (March 2023).**

FERRARI A. and V. NISPI LANDI, *Whatever it takes to save the planet? Central banks and unconventional green policy,* Macroeconomic Dynamics, v. 28, 2, pp. 299-324, **WP 1320 (February 2021).**

FLACCADORO M., *Exchange rate pass-through in small, open, commodity-exporting economies: lessons from Canada,* Journal of International Economics, v. 148, Article 103885, **WP 1368 (May 2022).**

MORO A. and V. NISPI LANDI, *The external financial spillovers of CBDCs,* Journal of Economic Dynamics and Control, v. 159, Article 104801, **WP 1416 (July 2023).**

FORTHCOMING

BALTRUNAITE A., M. CANNELLA, S. MOCETTI and G. ROMA, *Board composition and performance of state-owned enterprises: quasi experimental evidence,* The Journal of Law, Economics, and Organization, **WP 1328 (April 2021).**

BENVENUTI M. and S. DEL PRETE, *The evolution of banking competition in italian credit markets using a profit elasticity approach,* Italian Economic Journal / Rivista italiana degli economisti, **WP 1237 (October 2019).**

CIANI E., A. GROMPONE and E. OLIVIERI, *Jobs for the long-term unemployed: place-based policies in depressed areas,* Italian Economic Journal / Rivista italiana degli economisti, **WP 1249 (November 2019).**

CUCINIELLO V. and N. DI IASIO, *Determinants of the credit cycle: a flow analysis of the extensive margin,* Journal of Money, Credit and Banking, **WP 1266 (March 2020).**

FLACCADORO M., *Exchange rate pass-through in small, open, commodity-exporting economies: lessons from Canada,* Journal of International Economics, **WP 1365 (April 2022).**

GAUTIER E., C. CONFLITTI, R. FABER, B. FABO, L. FADEJEVA, V. JOUVANCEAU, J.-O. MENZ, T. MESSNER, P. PETROULAS, P. ROLDAN-BLANCO, F. RUMLER, S. SANTORO, E. WIELAND and H. ZIMMER, *New facts on consumer price rigidity in the euro area,* American Economic Journal: Macroeconomics, **WP 1375 (July 2022).**

LOSCHIAVO D., F. TULLIO and A. DI SALVATORE, *Measuring households' financial fragilities: an analysis at the intersection of income, financial wealth, and debt,* Review of Income and Wealth, **WP 1452 (April 2024).**

MICHELANGELI V. and E. VIVIANO, *Can internet banking affect households' participation in financial markets and financial awarness?,* Journal of Money, Credit and Banking, **WP 1329 (April 2021).**

MISTRETTA A., *Synchronization vs transmission: the effect of the German slowdown on the Italian business cycle,* International Journal of Central Banking, **WP 1346 (October 2021).**

MODENA F., S. PEREDA-FERNANDEZ and G. M. TANZI, *On the design of grant assignment rules,* Politica economica/Journal of Economic Policy, **WP 1307 (December 2020).**

MORO A., *Optimal policies in a small open economy with an environmental externality and shallow foreign exchange markets,* Portuguese Economic Journal, **WP 1348 (October 2021).**

RAINONE E., *Reservation rates in interbank money markets,* Journal of Money, Credit and Banking, **WP 1160 (February 2021).**

RAINONE E., *Real-time identification and high frequency analysis of deposits outflows,* Journal of Financial Econometrics, **WP 1319 (December 2017).**

ROPELE T., Y. GORODNICHENKO and O. COIBION, *Inflation expectations and misallocation of resources: evidence from Italy,* American Economic Review: Insights, **WP 1437 (December 2023).**