EQUILIBRIUM

Quarterly Journal of Economics and Economic Policy
2017 VOLUME 12 ISSUE 3, September
p-ISSN 1689-765X, e-ISSN 2353-3293
www.economic-policy.pl

ORIGINAL PAPER

Citation: Mazur, B. (2017). Probabilistic predictive analysis of business cycle fluctuations in Polish economy. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 12*(3), 435–452. doi: 10.24136/eq.v12i3.23

Contact: blazej.mazur@uek.krakow.pl, Cracow University of Economics, Chair of Econometrics and Operations Research, Rakowicka 27, 31-510 Kraków, Poland

Received: 23 May 2017; Revised: 4 August 2017; Accepted: 15 August 2017

Błażej Mazur
Cracow University of Economics, Poland

Probabilistic predictive analysis of business cycle fluctuations in Polish economy

JEL Classification: *E37; C53*

Keywords: *density forecasts; indicator of future economic conditions; business cycle; Dynamic Conditional Score models; Generalized t distribution*

Abstract

Research background: The probabilistic setup and focus on evaluation of uncertainties and risks has become more widespread in modern empirical macroeconomics, including the analysis of business cycle fluctuations. Therefore, forecast-based indicators of future economic conditions should be constructed using density forecasts rather than point forecasts, as the former provide description of forecast uncertainty.

Purpose of the article: We discuss model-based probabilistic inference on business cycle fluctuations in Poland. In particular, we consider model comparison for probabilistic prediction of growth rates of the Polish industrial production. We also develop a class of indicators of future economic conditions constructed using probabilistic information on the rates (that make use of joint predictive distribution over several forecast horizons).

Methods: We use Bayesian methods (in order to capture the estimation uncertainty) and consider two groups of models. The first group consists of Dynamic Conditional Score models with the generalized t conditional distribution (with conditional heteroskedasticity and heavy tails, being important for modelling of extreme observations). Another group of models relies on deterministic cycle modelling using Flexible Fourier Form. Ex-post density forecasting performance of the models is compared using the criteria for probabilistic prediction: Log-Predictive Score (LPS) and Continuous Ranked Probability Score (CRPS).

Findings & Value added: The pre-2013 data support the deterministic cycle mod-els whereas more recent observations can be explained by a simple mean-reverting Gaussian AR(4) process. The results indicate a structural change affecting Polish business cycle fluctuations after 2013. Hence, forecast pooling strategies are recommended as a tool for further research. We find rather limited support in favor of the first group of models. The probabilistic indicator of future economic conditions considered here leads actual phases of the growth cycle quite well, though the effect is less obvious after 2013.

Introduction

The purpose of the paper is to set up a methodology that allows for practical predictive business cycle analysis based on industrial production data. We assume that inference about future evolution of business cycle conditions should be model-based and take into account the estimation and the prediction uncertainty. In other words, a model that is used to generate forecasts underlying any analysis of future business conditions should display satisfactory performance not only in terms of point forecasts, but also density forecasts. The density forecast is constructed as a joint (potentially multivariate over horizons) distribution which provides a formal description of uncertainty as to future values of the analyzed variable.

In the paper we make an effort to develop such a model, adequate for Polish data on industrial production. In order to do so we consider a menu of alternative specifications and discuss their properties as well as out-of-sample predictive performance. We make use of the forecasts to construct a probabilistic indicator describing future prospects as to the growth rates of the industrial production index. The indicator reflects forecast uncertainty as well as cross-horizon dependence. The concept of indicator is not necessarily the most widely used one, given, for example, an alternative formulation by Barhoumi *et al.* (2016).

The approach pursued here requires a number of problems to be addressed. Firstly, a univariate dynamic model for the industrial production series has to be constructed. The focus here is on business cycle properties and forecasting of the headline growth rate of industrial production (i.e. the year-on-year growth rate). However, it is not obvious whether it should be modelled directly (using year-on-year growth rates) or indirectly (using month-on-month growth rates). Within the indirect approach, one could model and forecast the month-on-month growth rates, and the forecast of year-on-year rate would be induced afterwards. The difference is of vital importance from the modelling point of view — the two processes display very different empirical properties. In particular, business-cycle-like fluctuations contribute a lot to the variance of the year-on-year rates, whereas for the month-on-month rates seasonal effects are the dominant ones.

Moreover, as the density forecasting perspective is taken here, the transformation from month-on-month growth rates to year-on-year growth rates relies on the whole multivariate distribution (the dimension is that of the forecast horizon). In other words, it is not possible to move between the two approaches considering the marginal (horizon-specific and univariate) density forecasts only. The cross-horizon stochastic dependence, which is not vital for point forecasts, turns out to be crucial for density forecasts. The dependence between forecasts for different horizons in the month-on-month setup has direct influence on dispersion of year-on-year forecasts in longer horizons. As a consequence, the availability of all the horizon-specific marginal information in one approach is not sufficient to recover the horizon-specific marginal information within the other approach. Therefore, a fan-chart of month-on-month rates is not enough to recover a fan-chart of year-on-year rates (and vice-versa).

The shift in attention from point forecasts towards the probabilistic (or density) forecasts is quite widespread in the recent econometric literature (see: e.g. Clark & Ravazzolo, 2015). The density perspective has been considered within the Bayesian approach for many years, since predictive distribution is a natural element of Bayesian inference. However, the application-oriented non-Bayesian econometric literature has given full appreciation to the probabilistic perspective in past few decades, in particular after the Global Financial Crisis.

The inspiration for the use of the probabilistic approach in empirical macroeconomics comes from the developments of statistical inference methods, sometimes related to some other applied areas e.g. weather forecasting (compare the discussion and the references cited by Lerch *et al*., 2017) or other branches of economics (see the analysis of energy markets by Nowotarski & Weron, 2017). Gneiting and Raftery (2007) have provided an influential discussion of state-of-the-art as to formal, statistical evaluation of density forecasts. The paper contains references to so-called proper scoring rules and strictly proper scoring rules that should be used for *ex-post* evaluation of density forecasts. Such criteria are log-predictive score (LPS) and continuous ranked probability score (CRPS) which are used in the empirical part below.

However, although the criteria mentioned above give full credit to the probabilistic perspective, a specific purpose of the paper might require additional considerations. This is because the formal statistical comparison measures goodness of fit that includes both short-term and long term behavior. However, for the purpose considered here i.e. the predictive analysis of business cycle fluctuations, the long-run adequacy is crucial. It might turn out that a model that has good properties in terms of capturing long-term properties of the data, but fails to capture short-term fluctuations, might rank rather low according to the formal criteria despite its usefulness for the purpose considered here. Hence the formal comparison might be augmented by less-formal and subjective *ex-post* evaluation of the forecasts.

Another related issue is that of structural change. It is not impossible that the underlying economic process driving the business cycle fluctuations is not time-homogenous (see: e.g. Bjørnland *et al*., 2017). Hence it is necessary to consider the problem of possible changes in adequacy of the competing models. In the case of Polish economy, a number of reasons support the view that some sort of structural change might have affected the dynamics of economic growth. For example it is not obvious that the pattern of business cycle fluctuations that has been identified for the Polish economy before, say, 2013 can be still considered adequate afterwards. The issue has serious consequences for the problem of model choice. Consequently, the predictive accuracy of competing specifications has to be evaluated in a dynamic way in order to identify possible shifts in forecasting performance caused by potential, underlying structural changes.

In the paper we focus on predictive adequacy, and we do not make an attempt to forecast industrial production in a multivariate setup. Instead, we focus on fine-tuning of a univariate specification in terms of more subtle properties like the form of the conditional distribution and dynamic evolution of conditional mean and variance — this is because long-term predictive properties are of interest here.

The rest of the paper is organized as follows. Firstly, we outline the classes of models used in the paper to forecast dynamics of Polish industrial production. The models include ones that emphasize business-cycle-like fluctuations (a deterministic cycle models, see: Lenart *et al*., 2016, as well as Lenart & Mazur, 2017), though other aspects of dynamic behavior are not modelled in a very sophisticated manner. Alternatively, we consider specifications that contain weaker assumptions as to the cyclical behavior, but are equipped with more complicated stochastic features, like heavy-tailed conditional distribution or time-varying conditional variance. The models are applied to the industrial production data, and their *ex-post* forecasting performance is thoroughly examined in an expending-window experiment with full recursive estimation. The question of interest is that of tracing changes in forecasting performance in the recent years. Finally, the best-performing models under consideration are used to generate an indicator that reflects the probabilistic information about future changes of growth rates of the industrial production index.

Research method

The models used here can be divided into two groups: the ones that explicitly account for the cyclical properties of the data, using the deterministic cycle idea (hence generating strong out-of-sample results) and the ones that rely on more sophisticated stochastic properties (in order to avoid failures in probabilistic predictive ability caused by too trivial stochastic formulation). The models of the first group make use of so-called Flexible Fourier Form in order to capture out-of-sample business cycle fluctuations, relying on a simple autoregressive formula with conditionally Gaussian observations. Models of the other group follow the idea of Dynamic Conditional Score (DCS) approach of Harvey (2013) and make use of more flexible conditional distribution, (namely the generalized *t* distribution). The basic structure of the models is recalled below, with references providing a more detailed description. The generalized *t* distribution used here is also briefly characterized. Some details regarding Bayesian model specification and estimation are provided as well. The model comparison in the empirical of the paper relies on evaluation criteria for density forecasts that are also summarized here.

Business cycle fluctuations are often modelled using Markov Switching models (see: e.g. Billio *et al*., 2016 or Eo & Kim, 2016), Dynamic Factor Models (e.g. Barhoumi *et al*., 2016) or non-linear models (Ferrara *et al*., 2016). However, the results obtained so far for the Polish economy indicate the relevance of deterministic cycle models, hence the model choice in the paper is application-specific.

The Bayesian model for analysis of deterministic cycle used here is discussed by Lenart & Mazur (2016, 2017), its application for in-sample business cycle analysis is considered by Lenart *et al*. (2016). The underlying idea is close to that of cyclostationarity: the mean of the process under consideration is time-varying, and the time-variation pattern is approximated using the Flexible Fourier Form (see: Gallant, 1981). It is assumed that short-term deviations from the time-varying mean μ_t are represented by a Gaussian autoregressive process denoted by v_t:

$$y_t = \mu_t + v_t,\ v_t = \psi_1 v_{t-1} + \cdots + \psi_p v_{t-p} + \varepsilon_t,\, \varepsilon_t \sim iiN(0, \sigma^2)$$

where y_t represents the observed series of year-on-year growth rates. The (Flexible Fourier) time-varying mean is given by:

$$\mu_t = \sum_{f=1}^{F} \left(\alpha_{1,f} \sin\left(t\phi_f\right) + \alpha_{2,f} \cos\left(t\phi_f\right)\right).$$

The parameters denoted by $\phi_f \in (\phi_L, \phi_U) \subseteq (0, \pi)$ represent frequencies of the fluctuations and the fixed lower and upper bounds (denoted by ϕ_L and ϕ_U) can be used in order to restrict attention to cyclical fluctuations of specific period length (i.e. to exclude fluctuations with period that is either too long or too short). The flexibility of the cyclical part of the model depends on the number of Fourier components (denoted by F). An interesting feature of the model is that it allows for $F > 1$, which implies that the business cycle fluctuations are driven by components with more than one empirically important frequency (which is rare in stochastic cycle models). In practical applications F is often restricted not to exceed 3, as higher values might lead to overfitting issues. Statistical inference in such model within the Bayesian setup is described in detail by Lenart and Mazur (2016). It is possible to generalize the approach into higher dimensions for e.g. cross-country analysis of business cycle synchronization (in order to obtain results similar in spirit to those of Lenart & Pipień, 2017), however this is left for further research.

The generalized t distribution used here was described by Theodossiou (1998), see also Theodossiou & Savva (2016). Its probability density function has the following form:

$$p(y) = \frac{1}{\sigma} K(\nu, \gamma) \left[1 + \frac{1}{\nu} \left(\frac{y - \mu}{\sigma} \right)^{\gamma} \right]^{-\frac{(1+\nu)}{\gamma}}$$

with:

$$K(\nu, \gamma) = \frac{\gamma}{2\nu^{1/\gamma}} \frac{1}{B\left(\frac{\nu}{\gamma}, \frac{1}{\gamma} \right)}.$$

The distribution given above has location parameter μ, scale parameter σ and two shape parameters: ν and γ. An interesting feature of this symmetric probability distribution is that it allows for heavy tails and encompasses a number of known distributions as nested or limiting cases. For example, as $\gamma = 2$, it becomes Student-t with ν degrees of freedom. On the other hand, with $\nu \to \infty$, the limiting case is GED(γ), the generalized error distribution; see Harvey and Lange (2017) for a more detailed discussion. The distribution is quite flexible and therefore capable of capturing many empirically relevant situations, especially related with the occurrence of rare events. The feature might be important for density predictive performance.

The Dynamic Conditional Score (DCS) models are discussed in detail by Harvey (2013). The model class is closely related to Generalized Autoregressive Score models of Creal *et al.* (2013), for a predictive application see e.g. Bernardi & Catania (2016). In the paper we follow the formulation by Harvey. However, our contribution is in developing methods of Bayesian inference for the models. Harvey makes use of the Maximum Likelihood estimation. The Bayesian model specification and inference is particularly important here, since the emphasis is on properties of the density forecasts obtained from the models. Within the maximum likelihood approach it is very difficult to derive density forecasts that take into account the estimation uncertainty. However, the uncertainty might be crucial especially for parameters that control more sophisticated properties of the distribution (like ν and γ that influence tail thickness in the case under consideration). On the other hand, within the Bayesian approach the estimation uncertainty is handled in a very natural way, as the predictive distribution is a mixture with mixing distribution being the full posterior for the model parameters.

The structure of Dynamic Conditional Score modelling reflects the idea that for a given conditional distribution, some of its features can be dynamically updated. The models are not based on latent stochastic processes, and well-known GARCH models represent similar reasoning, which leads to dynamically evolving conditional variance. However, in the case of DCS models the updating mechanism explicitly depends on score of the conditional distribution used (i.e. partial derivative of the log-density w.r.t. to the parameter under consideration). In other words, properties of the updating mechanism depend on the properties of the conditional distribution, which is a very appealing concept. The general idea together with numerous applications is described by Harvey (2013). Here we assume that the following formulation holds:

$$g_t = \eta_1 g_{t-1} + \cdots + \eta_p g_{t-p} + \phi_1 s_{t-1} + \cdots + \phi_q s_{t-q}$$

where g_t represents the deviation of the feature under consideration from its average (or seasonally changing) state: $g_t = f_t - \delta_t$, with δ_t being either time invariant ($\delta_t = \delta$) or seasonal ($\delta_t = \delta_{s(t)}$) with the initial conditions described by $g_0 \ldots\ g_{-p+1}$. Moreover, s_t is the value of the score at the point corresponding to the realized observation ad time *t*. This setup allows for e.g. seasonal effects in conditional volatility similar in the spirit to the approach of Lenart (2017).

We assume that the feature being updated (f_t) corresponds to the conditional location or the conditional scale. These represent the respective parameters of the Generalized *t* distribution introduced above. Consequently, the model is not formulated in terms of conditional moments, though due to symmetry of the distribution the relationship between scale and variance is not that complicated. For the scale parameter it is necessary to add so-called linking function that maps it values into the real line. Consequently, the linear autoregressive updating mechanism is applied to log-scale instead of scale. The additional linking function (logarithmic transformation) is taken into account when computing the score. When more than one feature is being updated, it is possible to consider a matrix version of the dynamic updating equation. However, the path is not pursued here. We assume that the updating mechanism is diagonal, i.e. works separately for each feature (however, there exists a relationship between the expression for score for the scale and the location parameters, see Harvey & Lange, 2017).

Within the above setup it is possible to consider seasonal and autoregressive dynamic effects in the conditional location or scale. For the purpose of the paper, the seasonality might be important if the variable being modelled represents month-on-month growth rates.

Bayesian specification and estimation of the above models is non-trivial. In particular one has to specify prior assumptions as to the model parameters and construct a working sampler that allows for exploration of the model posterior distribution. In all the models under consideration informative priors are imposed, with independence among groups of parameters. Important prior information is that of cycle length for the models with deterministic cycle. Estimation of the deterministic cycle models is undertaken using a hybrid Gibbs sampling scheme, which seems to be very efficient. However, for the dynamic conditional score models the estimation issue seems to be a serious one. The models are estimated using a Metropolis-Hastings algorithm (with independent proposal, being convenient for the sake of recursive prediction), though it is possible that some other method could display somewhat better mixing properties. Therefore there might be a room for improvement as to numerical performance for the Bayesian DCS models used here. A better posterior sampler might result in better approximation of the posterior and predictive densities.

The standard criteria for *ex-post* evaluation of the forecasts include RMSFE and MAE. However, the criteria are relevant to point forecasts only, hence convey some information about adequacy of the location of the density forecast, but completely ignore its dispersion (or other features of the distribution). However, from the decision-making point of view it is quite obvious that such a strong information reduction might be innocuous under very special conditions only. Here we assume that it is necessary to include other criteria for forecast evaluation as well.

Log-predictive score (LPS) used for *ex-post* evaluation requires computation of log-density value of predictive distribution at the actual outturn. It can be shown that LPS computed recursively for one-step-ahead forecasts is linked closely to some basic Bayesian measures of goodness of fit, which provides additional theoretical justification. In practice computation of the LPS for the tail outcomes might be numerically challenging. Moreover, it might turn out that a model that is well specified in terms of conditional location but mis-specified in terms of, say, conditional variance or tail thickness, might achieve very poor scores based on LPS.

Another measure under consideration is Continuous Ranked Probability Score (CRPS). It can be perceived as a generalization of the absolute error (AE), since if one assumes that the forecast distribution is point mass, CRPS is equal to AE. The measure is less sensitive to tail outcomes (compared to LPS) and poses no serious numerical challenges. Theoretical foundations underlying LPS and CRPS are discussed by Gneiting & Raftery (2007).

Finally, we propose an indicator of future business conditions (a Future Business Conditions Indicator, FBCI) that relies on full predictive distribution of year-on-year growth rates. It is intended to be evaluated using monthly data (in order to keep the inflow of new information). However, the y-on-y growth rates at monthly frequency often display considerable short-term variation (even in the case of calendar-adjusted data). We therefore assume that the indicator (denoted by FTI_M, for Future Tendency Indicator) represents the probability that the average growth rate for the period covering e.g. t+4, t+5, t+6 is greater than the average growth rate for t+1, t+2 and t+3 (taking into consideration a three-month basis period, so M = 3). In other words, it measures the probability of the general positive tendency during the next 2M periods (here: six months), on average. Alternatively, it could be computed with e.g. M = 6 or M = 12 (instead of M = 3). However, it must be emphasized that the indicator does not convey information as to the magnitude of the growth, dealing just with the direction of change in the growth rate dynamics. However, it might be interpreted as reflecting future prospects as to the growth cycle, taking into account both the prediction uncertainty and the stochastic dependence between forecasts for various horizons, which is not usual for such indicators.

Empirical analysis of Polish industrial production data

The dataset under consideration i.e. year-to-year growth rates of Polish industrial production (in per-cents, monthly data, adjusted for calendar effects, not seasonally adjusted, 1997M01-2016M12, $T = 240$) is depicted in Figure 1. We treat the first 120 observations (10 years) as a training sample, and verify the out-of-sample predictive performance of alternative models on the remaining 120 data points (full re-estimation is conducted with each observation added). The forecasts are generated within the expanding-window setup. No effort is made to mimic real-time data flow, instead the most recent readouts available are used. In what follows, we consider only the direct forecasts of y-o-y growth rates. This is because a preliminary analysis using various DCS models within the indirect approach indicates that such specifications deliver rather trivial (i.e. practically constant) forecasts of y-o-y rates for horizons greater than 12 months.

Consequently, in what follows we consider the models estimated on y-o-y data only and use the direct approach for the sake of prediction. We make use of two Gaussian autoregressive models, one with 4 lags and one with 22 lags (labeled AR(4) and AR(22)). The models are chosen to represent different degree of potential complexity of the autocorrelation function. Moreover, we consider a deterministic cycle model, with $F = 3$, frequency parameters restricted to the (0.052, 0.52) interval and 22 lags in the autoregressive part (labeled AR(22)-F(3)). The last specification under consideration is a DCS model with $p = q = 6$ for the location parameter and $p = q = 2$ for the log-scale parameter. The model allows for asymmetric response to the score (following Harvey & Lange, 2017) and its conditional distribution is of the generalized t form (Gt-DCS(6,6;2,2), labeled DCS for short).

Table 1 contains characteristics summarizing *ex-post* properties of point and density forecasts obtained from the models mentioned above.

The results are reported for horizons of 12, 18 and 24 months ahead, and also calculated using the last 36 realized forecasts only (the last observation used for the purpose of evaluation is that representing 2017M01). Analysis of Table 1 seems to lead to a very simple conclusion. The overall predictive performance is dominated by the Gaussian AR(22) model, and if one restricts attention to the last 3 years, the results support a simple AR(4) model. In particular, neither the stochastically sophisticated DCS specification nor the business-cycle oriented AR(22)-F(3) lead to satisfactory results. The conclusion is unanimously supported by point and density criteria.

However, a more detailed analysis can be conducted based on a decomposition of differences in cumulated LPS between certain models into the contribution of individual (realized) observations throughout the verification period. Such a decomposition for the two winning models (AR(22) and AR(4)) against the AR(22)-F(3) specification is presented in Figure 2 (A and B). The figures reveal the fact that throughout most of the verification window the data provide strong and systematic support in favor of the deterministic cycle model AR(22)-F(3). However, in 2013 the pattern breaks down and the predictive performance of the model deteriorates quickly (especially in 2014). As a consequence, the last three years of the data bring strong and prevailing evidence against the deterministic cycle model, and the empirical support shifts towards the AR(4) specification.

The abrupt change might suggest that that the business cycle properties of the Polish industrial production growth cycle have changed after 2013 in such way that the previously observed pattern (matching the deterministic cycle dynamics) was no longer valid. Importantly, the model with the best forecasting performance in the recent period, namely the AR(4), generates quite trivial forecasts: quick mean reversion results in rather flat forecast paths stabilizing at the sample mean. Lenart *et al*. (2016) analyze the following problem: has the pattern of Polish business cycle fluctuations changed after the Global Financial Crisis? Their conclusion, obtained using different methodological approaches, though based on somewhat shorter series and a sequence of recursive in-sample analyses, is negative — meaning that there is no evidence in favor of such a change. The results presented here contribute to the discussion indicating that the change might have occurred five years later, i.e. around 2013. Such a conclusion is, however, conditional upon specific assumptions used here, including the model set in particular. It is possible that the use of more advanced time-varying parameter models or dynamic prediction pooling strategies could shed more light on the problem. However, it seems clear that the deterministic cycle model of the Polish industrial production dynamics was adequate until 2013 only. It is though unclear what kind of model would be adequate for long-term prediction after 2013 — the issue could be considered as more data would be available.

The DCS-type models considered here are useful for prediction of month-on-month growth rates, though the implied (indirect) or direct year-on-year forecasts are not satisfactory, especially in longer horizons.

Finally, in Figure 3 we present the values of the probabilistic indicator of future economic conditions. Based on the above comparison, we pick the results for AR(4) and AR(22)-F(3). Moreover, the version presented here refers to one-year-ahead forecasts (M = 6). At each point in time the value of the indicator represents the probability that the average growth rate for t+7,…,t+12 will exceed its counterpart compared for t+1,…,t+6. Hence, the indicator provides information on the general direction of change (or the prevailing tendency) over the forecast horizon under consideration. Values greater than 0.5 indicate the prevalence of a positive trend during the next

2*M* periods. Conversely, values lower than 0.5 indicate the prevalence of a negative trend in the timespan ranging from 1 till 2*M* periods ahead.

One might notice that for most of the time the probabilistic indicator of future economic conditions obtained from AR(22)-F(3) provides clear-cut signals, being close to either 0 or 1, while the signals from the AR(4) model are somewhat less evident, though maintaining the same direction. Moreover, the indicator seems to provide adequate information (i.e. it leads the actual changes), at least in the first part of the sample. Closer to the sample end the fluctuations of industrial production growth rates dampen, so the adequacy of the indicator is more difficult to verify. This is also reflected by the values of the indicator, being close to 0.5 in the final part of the sample in the case of AR(4) model. In the case of AR(22)-F(3) model, the values close to the sample end still indicate rather positive prospects.

Conclusions

In the paper we compare density predictive performance of alternative model specifications with application to y-on-y growth rates of Polish industrial production. The objective of the research is to capture the out-of-sample business cycle fluctuations (in the form of the growth cycle). We consider two model classes. Specifications of the first group capture business cycle dynamics using the deterministic cycle approach based on Flexible Fourier Form (see Lenart & Mazur, 2016). Those of the second kind are more general in terms of the stochastic specification. The Dynamic Conditional Score models used here allow for heavy-tailed conditional distribution (of the generalized *t* class) and time-varying conditional scale.

We generate density forecasts for horizons up to 24 months ahead. Evaluation of the forecasts (based on CRPS and LPS criteria) seems to indicate that the DCS-type models do not generate additional predictive power despite their relative complexity. Closer examination reveals the fact that up to 2013 the best-performing model was that of deterministic cycle, while in more recent period the *ex-post* evidence shifts toward a simple Gaussian AR(4) specification with quick mean reversion. The results indicate a structural change in the process underlying Polish business cycle fluctuations. It is interesting that the change identified here does not seem to be related with the Global Financial Crisis of 2007/8 (as one could expect, see e.g. the analysis of Dąbrowski *et al*., 2015). As there seems to be clear evidence for model instability in the recent years, a practical suggestion to address this issue (and a direction for future research) is the use of dynamic pools of density forecasts for the sake of probabilistic forecasting of Polish industrial production series.

Moreover, we demonstrate a forward-looking (i.e. fully based on predictive results) probabilistic indicator of future economic conditions (labeled FTI_M). Its use is illustrated with an analysis of growth rates of the Polish industrial production. The values of the indicator correspond to the overall tendency prevailing in the whole forecast period. The indicator has two important features. Firstly, it is constructed using density forecasts, hence it takes the forecast uncertainty into account (here we make use of Bayesian methods, so the estimation uncertainty is also accounted for). Secondly, the indicator is based on joint predictive distribution over a sequence of horizons. It therefore goes beyond the horizon-specific information, utilizing the cross-horizon stochastic dependence as well. The latter feature is less tangible, as it is not reflected in usual fan-charts. However, the indicator does not provide information as to the magnitude of the predicted growth or decline. This might suggest a bivariate extension, providing two signals at the same time, reflecting probability of a change of given magnitude over a range of magnitudes.

References

Barhoumi, K., Darné, O., & Ferrara, L. (2016). A world trade leading index (WTLI). *Economics Letters*, *146*. doi: 10.1016/j.econlet.2016.07.032.

Bernardi, M., & Catania, L. (2016). Comparison of value-at-risk models using the MCS approach. *Computational Statistics*, *31*(2). doi: 10.1007/s00180-016-0646-6.

Billio, M., Casarin, R., Ravazzolo, F., & Van Dijk, H. K. (2016). Interconnections between Eurozone and US booms and busts using a Bayesian panel Markov-Switching VAR model. *Journal of Applied Econometrics*, *31*(7). doi: 10.1002/jae.2501.

Bjørnland, H. C., Ravazzolo, F., & Thorsrud, L. A. (2017). Forecasting GDP with global components: this time is different. *International Journal of Forecasting*, *33*(1). doi: 10.1016/j.ijforecast.2016.02.004.

Clark, T. E., Ravazzolo F. (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, *30*(4). doi: 10.1002/jae.2379.

Creal, D., Koopman, S. J. & Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, *28*. doi: 10.1002/jae.1279.

Dąbrowski, M. A., Śmiech, S., & Papież, M. (2015). Monetary policy options for mitigating the impact of the global financial crisis on emerging market economies. *Journal of International Money and Finance, 51*. doi: 10.1016/j.jimonfin.2014.12.006.

Eo, Y., & Kim, C.-J. (2016). Markov-switching models with evolving regime-specific parameters: are postwar booms or recessions all alike? *Review of Economics and Statistics, 98*(5). doi: 10.1162/rest_a_00561.

Ferrara, L., Marcellino, M., & Mogliani, M. (2015). Macroeconomic forecasting during the great recession: the return of non-linearity? *International Journal of Forecasting*, *31*(3). doi: 10.1016/j.ijforecast.2014.11.005.

Gallant A. R. (1981). On the bias in flexible functional forms and an essentially unbiased form: the Fourier flexible form. *Journal of Econometrics, 15*. doi: 10.1016/0304-4076(81)90115-9.

Gneiting, T., & Raftery, A. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, *102*(477). doi: 10.1198/016214506000001437.

Harvey, A. C. (2013). *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*. Cambridge: Cambridge University Press. doi: 10.1017/cbo9781139540933.

Harvey, A. C., & Lange, J.-R. (2017). Volatility modelling with a generalized *t*-distribution. *Journal of Time Series Analysis, 38*. doi: 10.1111/jtsa.12224.

Lenart, Ł. (2017). Examination of seasonal volatility in HICP for Baltic region countries: non-parametric test versus forecasting experiment. *Central European Journal of Economic Modelling and Econometrics, 9*(1).

Lenart, Ł., & Mazur, B. (2016). On Bayesian Inference for Almost Periodic in Mean Autoregressive Models. *Przegląd Statystyczny*, *63*(3).

Lenart, Ł., & Mazur, B. (2017). Business cycle analysis with short time series: a stochastic versus a non-stochastic approach. In M. Papież and S. Śmiech (Eds.), *The 11th Professor Aleksander Zelias international conference on modelling and forecasting of socio-economic phenomena. Conference proceedings.* Cracow: Foundation of the Cracow University of Economics. Retrieved form http://pliki.konferencjazakopianska.pl/proceedings_2017/.

Lenart, Ł., Mazur, B., & Pipień, M. (2016). Statistical analysis of business cycle fluctuations in Poland before and after the crisis. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *11*(4). doi: 10.12775/equil.2016.035.

Lenart, Ł., & Pipień, M. (2017). Non-parametric test for the existence of the common deterministic cycle: the case of the selected European countries. *Central European Journal of Economic Modelling and Econometrics*, *9*(3).

Lerch, S., Thorarinsdottir, T. L., Ravazzolo, F., & Gneiting, T. (2017). Forecaster's dilemma: extreme events and forecast evaluation. *Statistical Science*, *32*(1). doi: 10.1214/16-sts588.

Nowotarski, J., & Weron, R. (2017). Recent advances in electricity price forecasting: A review of probabilistic forecasting. *Renewable and Sustainable Energy Reviews*, doi: 10.1016/j.rser.2017.05.234.

Theodossiou, P. (1998). Financial data and the skewed generalized *t* distribution. *Management Science*, *44*. doi: 10.1287/mnsc.44.12.1650.
Theodossiou, P., & Savva, C. S. (2016). Skewness and the relation between risk and return. *Management Science*, *62*(6). doi: 10.1287/mnsc.2015.2201.

Acknowledgments

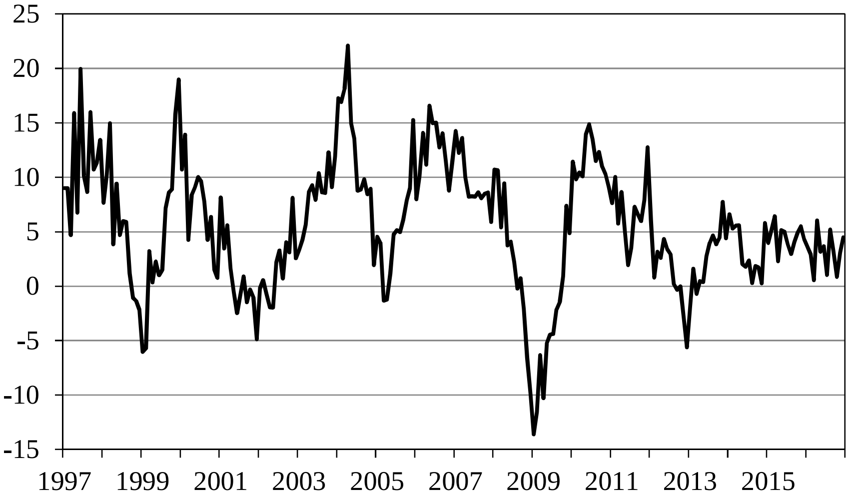
This research was supported by the Polish National Science Center based on decision number DEC-2013/09/B/HS4/01945.

Annex

Table 1. *Ex-post* evaluation of point and density forecasts (12, 18 and 24 months ahead). LPS is computed using natural logs (cumulated, the higher the better), CRPS is in positive orientation (averaged, the lower the better)

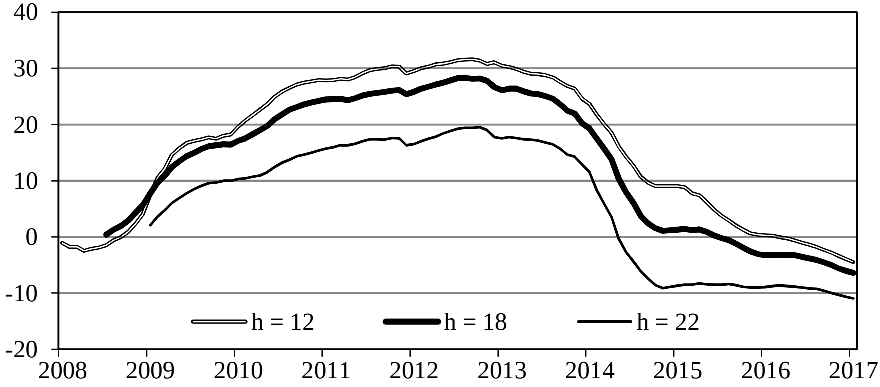
	h = 12			*h* = 18			*h* = 24		
	RMSE	LPS	CRPS	RMSE	LPS	CRPS	RMSE	LPS	CRPS
full verification window: (120 – *h*) observations									
AR(4)	6.46	-359.47	3.52	6.50	-340.84	3.58	5.58	-314.52	3.30
AR(22)	5.83	-351.57	3.30	5.87	-332.48	3.32	5.12	-308.72	3.10
AR(22)-F(3)	6.72	-363.96	3.96	6.80	-347.25	3.97	6.36	-325.44	3.79
DCS	6.77	-370.00	3.82	6.83	-350.14	3.85	6.13	-323.79	3.61
last 36 observations									
AR(4)	2.67	-103.44	1.94	2.48	-104.77	1.95	2.73	-106.05	2.06
AR(22)	4.79	-109.62	2.82	4.53	-109.01	2.69	3.81	-107.66	2.39
AR(22)-F(3)	8.14	-131.54	5.02	7.81	-130.48	4.73	7.08	-128.48	4.21
DCS	6.08	-114.43	3.57	5.93	-113.68	3.35	4.93	-110.70	2.85

Figure 1. Growth rates of Polish industrial production index (y-o-y, in [%])

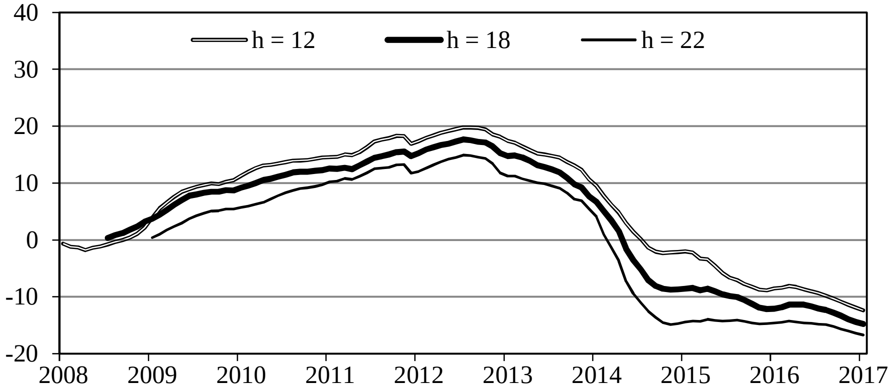


Source: Eurostat.

Figure 2. Differences in cumulated LPS between AR(22)-F(3) and other models, recursive expanding-window estimation, positive values support AR(22)-F(3)



A: differences vs. AR(4)



B: differences vs. AR(22)

Figure 3. Values of the FTI_6 indicator obtained recursively in an expanding-sample setup based on sequences of forecasts from two models, AR(22)-F(3) and AR(4) (left axis) and the actual data (right axis)