F O L I A   O E C O N O M I C A   C R A C O V I E N S I A
Vol. LVIII (2017) PL ISSN 0071-674X

PROBABILISTIC PREDICTION USING DISAGGREGATE DATA: THE CASE OF GROSS VALUE ADDED IN POLAND[1]

BŁAŻEJ MAZUR
Cracow University of Economics, Department of Econometrics and Operations Research
e-mail: mazurb@uek.krakow.pl

STRESZCZENIE

B. Mazur. *Predykcja wartości dodanej brutto w polskiej gospodarce: ujęcie probabilistyczne z wykorzystaniem danych zdezagregowanych.* Folia Oeconomica Cracoviensia 2017, 58: 85–103.

W artykule przeprowadzono analizę *ex-post* zdolności predyktywnych (w odniesieniu do prognoz probabilistycznych) alternatywnych specyfikacji, odzwierciedlających występowanie regularnych wahań koniunkturalnych. Celem pracy jest stwierdzenie czy modelowanie danych zdezagregowanych pozwala uzyskać lepsze prognozy probabilistyczne wielkości zagregowanej (w porównaniu z podejściem, w którym modelowaniu podlega bezpośrednio sam agregat). Empiryczne porównania ujęcia zagregowanego i zdezagregowanego są przedstawione w literaturze głównie dla prognoz punktowych. W pracy wykorzystano wnioskowanie bayesowskie i dokonano porównania rozkładów prognoz agregatu, indukowanych na podstawie rozkładów predyktywnych z modeli wielowymiarowych dla subagregatów (ze swoistymi komponentami cyklicznymi dla każdej składowej) oraz rozkładów prognoz dla agregatu otrzymanych z modelu jednowymiarowego z komponentem cyklicznym. Wielkością prognozowaną jest dynamika r/r wartości dodanej brutto w polskiej gospodarce (w cenach stałych), eksperyment predyktywny uwzględnia publikowane rewizje danych. Otrzymane wyniki wskazują, że w ramach ujęcia zdezagregowanego udało się uzyskać rozkłady prognoz o lepszych (*ex post*) własnościach.

ABSTRACT

The objective of the paper is to compare *ex-post* performance of density forecasts based on alternative specifications that allow for regular business-cycle type fluctuations. The empirical question posed here is whether the disaggregate (i.e. indirect) approach to forecasting of the aggregate quantity is capable of delivering results that are superior to forecasts obtained in a direct way (i.e. applying a univariate model to the aggregate quantity). Such comparisons have been conducted for point forecasts, but the literature dealing with density forecasts is still scarce. We make use of the Bayesian approach. The predictive distribution of the aggregate quantity can

[1] The research was supported by the funds granted to the Faculty of Management, Cracow University of Economics within the framework of the subsidy for the maintenance of research potential.

be formally induced from a predictive distribution for the disaggregate quantities obtained in a multivariate model. The latter accounts for possible heterogeneity as it might include individual cyclical components for each variable. Alternatively, one might obtain forecasts of the aggregate quantity directly from a univariate model containing one cyclical component. The quantity of interest is the total gross value added in the Polish economy (with quarterly observations on real year-on-year growth rates). The predictive experiment conducted here takes into account data revisions, getting close to the real-time setup. Its results indicate that the forecasts obtained using models for disaggregate data have better *ex-post* properties.

SŁOWA KLUCZOWE — KEY WORDS

rozkłady prognoz, dane zdezagregowane, prognozowanie makroekonomiczne, wahania koniunkturalne, wnioskowanie bayesowskie

density forecasts, disaggregate data, macroeconomic forecasting, business cycles, Bayesian inference

1. MOTIVATION

Forecasting of business cycle fluctuations is of obvious practical importance nowadays. Moreover, the need for the use of the probabilistic approach in macroeconometrics is broadly recognized. In general, the attention has shifted towards the models that are capable of delivering satisfactory density forecasts, going beyond highlighting point forecasts as before. However, many practical approaches used for the purpose of macroeconomic forecasting have been developed based on point forecast performance evaluation. Hence, it is not obvious that such models would be successful in delivering density forecasts. Satisfactory performance of density forecasts relies on adequate modelling of e.g. *ex-ante* forecast dispersion, tail thickness etc. in addition to location.

Many studies suggest that including a stochastic volatility component improves density forecasting performance in the case of macroeconomic data, see e.g. Clark and Ravazzolo (2015), see also Mazur (2017) for a specific counterexample. The presence of a stochastic volatility component implies that the variance of predictive distribution can adjust to shocks in a more flexible way and allows for potential heavy tails in density forecasts. We aim at consideration of another important problem of macroeconomic density forecasting. There are numerous cases in which the quantity of interest allows for disaggregate modeling; see e.g. Lenart and Leszczyńska-Paczesna (2016) or Mazur (2015). In such a situation predictive variance of the aggregate quantity (implied from a multivariate model for sub-aggregates) depends strongly on assumptions on cross-variable contemporaneous dependence, which is at risk of mis-specification. In other words, certain model assumptions might be of small importance for properties of point forecasts of the aggregate quantity but essential for the purpose of density prediction.

Another issue of direct relevance addressed here is that of identification of cyclical pattern characterizing macroeconomic quantities of interest; see Harvey et al. (2007), Lenart et al. (2016), Lenart (2017), Lenart and Mazur (2017), Lenart and Pipień (2017). Macroeconomic analyses often deal with quite limited information content (i.e. short time series that are subject to revisions), as in the case of Poland. It is therefore easy to obtain weak or spurious results as to trend or cycle components of the series under consideration. Moreover, the aggregate quantity might be a function of sub-aggregates having heterogeneous properties (e.g. with certain sub-aggregates displaying clear and regular periodic pattern whilst some others being more stochastic and irregular). This is even more likely to be important in the case where shares of the sub-aggregates (in the aggregate quantity) change over time. The aim of the paper is to provide empirical evidence as to whether the business-cycle like behavior should be modelled at the aggregate or disaggregate level (for the sake of density forecasting of the aggregate quantity), see also Stock and Watson (2014).

It is also unclear how the optimal level or scope of disaggregation should be chosen. Fully formal techniques of model comparison are of no direct help. This is because there is no formal and generally accepted approach to the problem of goodness-of-fit comparison across different disaggregation spaces (i.e. across models of different dimensions with potentially overlapping information content). Therefore, it is necessary to make use of e.g. scoring rules to compare predictive performance of different models with respect to prediction of certain common (e.g. aggregate) quantity. However, such a comparison is not fully coherent. Firstly, many different criteria exist and the resulting model ranking can be sensitive to the choice. Secondly, it is difficult to develop so-called coherence criteria for prior specifications of different disaggregate models (this problem is specific to Bayesian inference). Consequently, the researchers interested in macroeconomic forecasting have to rely on empirical analyses that lack fully formal theoretical background. Therefore the purpose of the paper is to verify whether the problem is worth investigation at all from the empirical viewpoint. In other words the question addressed here is whether in a realistic empirical setup it is possible to develop a disaggregate model that delivers density forecasts of the aggregate quantity that are superior to forecasts obtained directly using univariate models.

Additional motivation for consideration of the approach based upon disaggregate data goes beyond the issues of pure statistical inference. As macroeconomic forecasts are often prepared to support decision-making process (or policy making), the ability to develop a coherent interpretation of the results is of great importance. A usual economic interpretation relies not only on e.g. forecasts of the GDP growth rate, but also on inference on sources of the economic growth (or decline). If an aggregate forecast is implied from disaggregate forecasts, it is possible to develop a coherent interpretation that includes not only inference on future dynamics of the aggregate quantity, but also provides insights into structure of the process. Contrary, if the disaggregate and the aggregate forecasts are obtained using separate models, it is very likely that the results might be incoherent (i.e. the disaggregate data would imply a different forecast of the aggregate quantity compared to that obtained from a univariate model). In such a situation in would be difficult to develop a convincing interpretation of the results that could be useful for policy making.

The rest of the paper is organized as follows. Firstly, the dataset under consideration is described, with explanation of alternative aggregation schemes and the treatment of data revisions. Secondly, the general model setup is introduced. The model class used here can be perceived as a generalization to univariate specifications described by Lenart and Mazur (2016). The resulting multivariate model class includes variable-specific Flexible Fourier-type fluctuations in mean with VAR-type deviations around it. Such (non-stationary in mean) setup allows for capturing of both heterogeneity and cross-variable dependence. Thirdly, the expanding sample quasi-real-time forecasting experiment is conducted. The *ex-post* empirical comparison of forecasting performance of alternative specifications is based on the use of scoring rules for density forecast evaluation (log-predictive score, LPS, and continuous ranked probability score, CRPS). Results of the comparison indicate that the 'indirect' density forecasts of the total gross value added (GVA) in Polish economy based on disaggregate data are superior to 'direct' forecasts obtained from univariate models for the aggregate quantity only, once the cross-series heterogeneity is accounted for.

2. THE DATA

Gross value added (GVA) for the Polish economy is being reported on quarterly basis by the Central Statistical Office of Poland. Since the third quarter of 2011 the CSO delivers the data upon ten components of the total GVA according to the disaggregation scheme described below (see Table 1).

Table 1

Components of the total GVA reported by the CSO (GUS)

	Gross Value Added: components
$Z_{(1)}$	Industry
$Z_{(2)}$	Construction
$Z_{(3)}$	Trade; repair of motor vehicles
$Z_{(4)}$	Transportation and storage

Table 1

	Gross Value Added: components
$Z_{(5)}$	Accommodation and catering
$Z_{(6)}$	Information and communication
$Z_{(7)}$	Financial and insurance activity
$Z_{(8)}$	Real estate activities
$Z_{(9)}$	Professional scientific and technical activities. Administrative and support service activities
$Z_{(10)}$	Public administration and defence; compulsory social security. Education. Human health and social work activities
$Z_{(11)}$	Other activities (implicitly defined)

Source: Central Statistical Office of Poland (GUS).

Observations following the disaggregation scheme are available back to 2002 (for levels), at quarterly frequency. One component (here labelled $Z_{(11)}$) is not reported but once perfect aggregation is assumed, its values can be computed based on values reported for the aggregate series. Unfortunately, the series are relatively short for an *ex-post* predictive experiment (with $T = 56$, as the first four observations are used to compute growth rates and the last observation available is 2017Q1, which is used for the *ex-post* verification). Consequently, it is practical to aggregate the series into a smaller number of manageable components, as the number of parameters capturing the cross-variable dependence increases quickly with the number of variables. In the paper we consider two disaggregation patterns, one with four components and one with five components. Two important components (representing 'Industry' and 'Trade, repair of motor vehicles') are common. Therefore we consider seven disaggregate series in total. Definitions of the variables are given in Table 2.

Table 2

Aggregation scheme used in the paper: definitions of variables $X_{(i)}$

		Gross value added (GVA): the components:
$X_{(1)}$	$Z_{(1)}$	Industry
$X_{(2)}$	$Z_{(3)}$	Trade; repair of motor vehicles
$X_{(3)}$	$Z_{(2)} + Z_{(4)} + Z_{(5)} + Z_{(7)}$	Construction, Transportation and storage, Accommodation and catering, Financial and insurance activity

Table 2

$X_{(4)}$	$Z_{(6)} + Z_{(8)} + Z_{(9)} + Z_{(10)} + Z_{(11)}$	Information and communication, Real estate activities, Professional scientific and technical activities. Administrative and support service activities, Public administration and defence; compulsory social security. Education. Human health and social work activities, Other activities
$X_{(5)}$	$Z_{(2)} + Z_{(7)} + Z_{(8)}$	Construction, Financial and insurance activity, Real estate activities
$X_{(6)}$	$Z_{(4)} + Z_{(5)} + Z_{(6)} + Z_{(9)}$	Transportation and storage, Accommodation and catering, Information and communication, Professional scientific and technical activities. Administrative and support service activities,
$X_{(7)}$	$Z_{(10)} + Z_{(11)}$	Public administration and defence; compulsory social security. Education. Human health and social work activities, Other activities

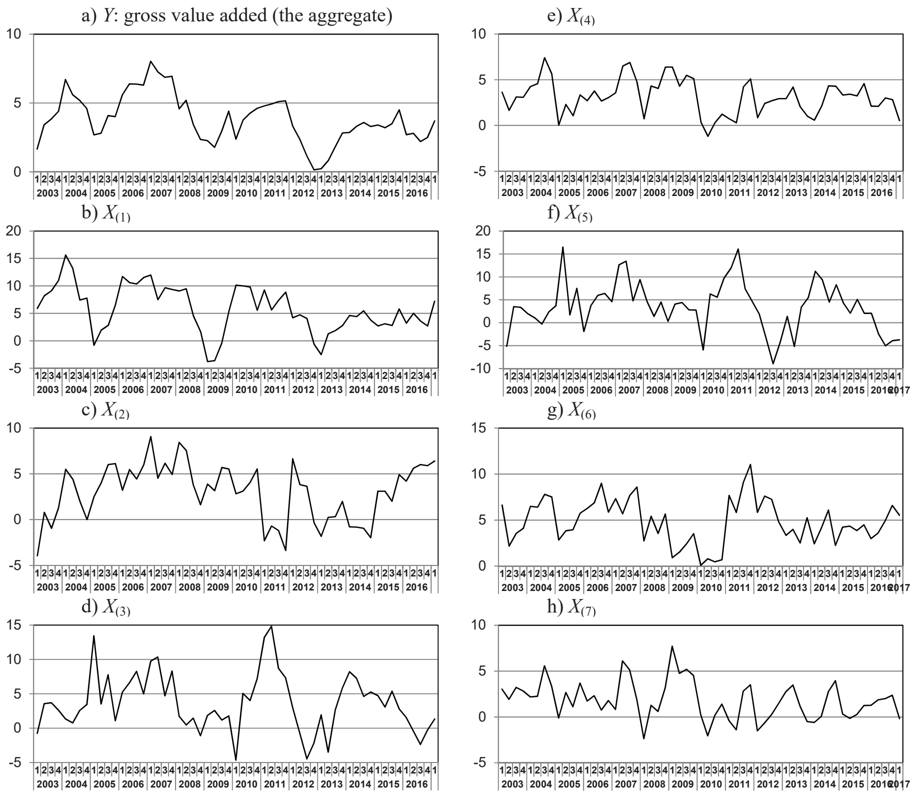
Source: Central Statistical Office of Poland (GUS).

In what follows we will consider multivariate models for two sets of variables: $\{X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}\}$ and $\{X_{(1)}, X_{(2)}, X_{(5)}, X_{(6)}, X_{(7)}\}$ representing alternative disaggregation patters of the total GVA (labelled Y). In general, we assume that the observed growth rates of y_t are denoted by $r_{t,y}$ (with $r_{t,y} = \left(y_t / y_{t-4} - 1\right) * 100\%$ for quarterly data). Hence the $X_{(i)}$'s are transformed into year-on-year growth rates, in per-cents. The growth rates (in constant prices) for $Z_{(1)} - Z_{(10)}$ and Y (not seasonally adjusted) are published by the CSO, together with aggregate monetary quantities that allow for computation of shares of $Z_{(i)}$'s in Y. Based on that it is possible to derive approximate shares and growth rates of $X_{(1)} - X_{(7)}$. Consequently, the GVA (Y) growth rates are given as:

$$r_{t,y} = \sum w_{j,t-4} r_{t,X(j)} \text{ for } t = 1, \ldots, T, \tag{1}$$

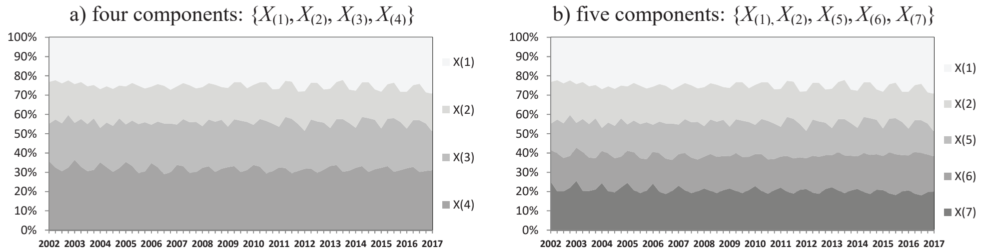
where the summation runs through the elements of $\{X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}\}$ or $\{X_{(1)}, X_{(2)}, X_{(5)}, X_{(6)}, X_{(7)}\}$ and $r_{t,Y}$ denotes growth rates of GVA (the aggregate quantity, Y) while $r_{t,X(j)}$ as well as $w_{j,t}$ represent growth rates and shares (in Y) of $X_{(j)}$ at time t. Therefore, based on forecasted values of $X_{(j)}$'s it is possible to predict the growth rates of Y. However, forecasts for horizons greater than four require that the shares w_j are also predicted. This can be done based on sample shares and predicted growth rates of $X_{(j)}$'s. In other words, within the disaggregate approach, prediction of growth rates of Y requires predicted values of $X_{(j)}$'s. Hence, a multivariate predictive model for $\{X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}\}$ or $\{X_{(1)}, X_{(2)}, X_{(5)}, X_{(6)}, X_{(7)}\}$ together with the observations on shares make it possible to induce the whole predictive distribution for Y. For the sake of probabilistic prediction of the aggregate growth rates, such a derivation is formally supported within the Bayesian approach; a similar reasoning could perhaps be conducted using other simulation-based methods like bootstrapping.

The variables under consideration are depicted in Figure 1, whereas the shares representing the two disaggregation patterns under consideration are shown in Figure 2. The data on national accounts are often revised or subject to methodological changes. The latter reason makes it difficult to find the relevant pre-2002 data. However, in order to take into account the revision process, we use a sequence of data vintages based on quarterly announcements of the CSO between 2011Q3 and 2017Q1. As the revisions published on the quarterly basis usually cover about 4 years of the data history, we assume that the data for the period 2002–2007 is fixed at values available at the time of the research (June 2017). Such a methodology results in availability of 23 vintages of the data approximating the real-time setup relatively well. We assume that the forecasts are evaluated using the latest outturn available (i.e. the most recent data vintage) which was used to construct the data depicted in Fig. 1 and 2.



Source: own computations using the Central Statistical Office of Poland (GUS) data.

Figure 1. The data: y-o-y growth rates [%] 2003:Q1–2017:Q1 (the latest vintage available)



Source: own computations using the Central Statistical Office of Poland (GUS) data.

Figure 2. Two disaggregation schemes: empirical shares in GVA (the most recent vintage), 2002Q1–2017Q1

Consequently, for one-quarter-ahead forecasts (practically equivalent to nowcasting) there are 22 realized values that can be used for *ex-post* evaluation of the predictive performance. The number of outturns available for longer horizons is obviously smaller, hence we restrict ourselves to analysis of $h = 1,\ldots,8$.

3. THE MODEL CLASS AND INFERENCE TECHNIQUE

The model framework used here expands the setup of Lenart and Mazur (2016) into many dimensions. The model makes use of the idea of Flexible Fourier Form to capture dynamic properties associated with business-cycle type of behavior. As the paper deals with dynamics of the growth rates (hence focusing on so-called growth cycle), the trends are not included into the model. Instead, the individual series display 'cyclical trends' modelled as in Mazur and Lenart (2016). Moreover, short-term deviations from the 'cyclical trends' are captured by a vector autoregressive structure, hence the cross-variable dependence (contemporaneous and lagged) is allowed for. The rationale for such a model is the following: the usual VAR-type (or BVAR-type) models are either stationary or based on so-called local level (i.e. an unobserved random walk driving the conditional mean of the process). Hence the resulting forecast paths are either stabilizing (mean reverting) or display a random-walk like behavior, approximately projecting the last observation into the future, with expanding uncertainty. The model considered here is non-stationary, so it does not generate forecast paths that converge (or stabilize at some fixed value). However, the non-stationarity is not of $I(1)$ type, hence the conditional (or predictive) variance does not explode. The model is an attempt to go beyond stationary vs. $I(1)$ alternatives that are most commonly used in the applied work.

Each observed quantity is modelled as being a sum of a cyclical component $\mu_{t,i}$ (intended to capture the long-run features i.e. business cycles) and a short-run deviation process labelled $u_{t,i}$:

$$r_{t,i} = \mu_{t,i} + u_{t,i} \text{ for } t = 1,\ldots,T; i = 1,\ldots,M. \tag{2}$$

The variable-specific cyclical component is represented by a Flexible Fourier Form-type construct; see Gallant (1981) and Lenart and Mazur (2016):

$$\mu_{t,i} = \gamma_i + \sum_{f=1}^{F_i} [\alpha_{f,i} \sin (t\phi_{f,i}) + \beta_{f,i} \cos (t\phi_{f,i})]. \tag{3}$$

The parameters of $\mu_{t,i}$ and $\mu_{t,j}$ ($j \neq i$) are not linked by any equality restrictions (so the cyclical components are fully variable-specific). It is assumed here that the short-run process of deviations u_t is a VAR(p) process:

$$u_t = \Theta_1 u_{t-1} + \cdots + \Theta_p u_{t-p} + \varepsilon_t, \tag{4}$$

with $\{\varepsilon_t\}$ denoting M-dimensional Gaussian white noise sequence with zero mean and time-invariant contemporaneous covariance matrix Ω. The assumption of conditional normality and homoscedasticity of disturbances ε_t might be considered too simplistic, though it is difficult to relax it with such a limited number of observations. In (3) the ϕ's control frequency (or cycle length) whereas α's and β's represent amplitudes and phase shifts. The form is deterministic, therefore assuming somewhat less flexible pattern of the fluctuations (see Lenart and Mazur (2017) for an alternative approach). Importantly, the above model allows for heterogeneity across the sub-aggregates with respect to their cyclical properties, as F_i's can take different values (zeros in particular). In general, higher values of F_i's provide increased flexibility, which comes at a risk of overfitting.

As to the frequency parameters (the ϕ's), the general assumption is that $\phi_{f,i} \in (0,\pi)$, however one might want to restrict the range of admissible values to an interval resulting in: $\phi_{f,i} \in (\phi_L, \phi_U) \subset (0,\pi)$. Setting known constants $0 < \phi_L < \phi_U < \pi$ amounts to excluding cycles with period lengths that exceed the limiting values. The relationship between period length ω_f and frequency ϕ_f is:

$$\omega_{f,i} = 2\pi / \phi_{f,i}, \tag{5}$$

with units of ω_f corresponding to the sampling frequency of r_t.

The model is estimated using the Bayesian approach, which requires setting priors for the model parameters. In particular the prior distribution for the frequency parameters ϕ_f is assumed to be iid uniform (though the implied prior for ω_f is not uniform, allocating less probability mass to lower frequencies or longer cycles). The prior imposes relatively weak structure, reducing the risk of identifying spurious low-frequency dynamics to some extent.

Prior distribution for the VAR parameters (Θ's) is of Minnesota type, though the values of scale hyperparameters are not assigned based on the data, but fixed at certain pre-determined values (hence the prior distribution actually does not depend upon the data). The model was estimated using a Gibbs sampler with M-H steps using author's own codes written in Ox; see Doornik (2007)[2]. The stationarity (non-explosive behavior) restrictions are not imposed upon the VAR parameters.

Alternative methodological approaches to business cycle modelling make use of subsampling methods — see e.g. Lenart and Pipień (2017). However, the subsampling methods are difficult to use in the case of short series (as the quarterly data from the Polish economy), whereas the above model works quite well within such a setup. Secondly, within the subsampling approach it is difficult to obtain density forecasts that account for the estimation uncertainty, which is true as to the Bayesian approach applied here.

4. EMPIRICAL ANALYSIS

The empirical results analyzed below are based on a set of recursive, expanding sample predictive experiments making use of 22 data vintages, with the shortest sample covering the period 2003:Q1–2012:Q3 while the longest subsample ends in 2016:Q4, with $T = 56$. All the models are fully re-estimated with each vintage (with at least 1 200 000 MCMC draws from the predictive distribution within each estimation/prediction step). Therefore, the setup quite closely mimics the real-time forecasting situation. We also assume that all the forecasts are evaluated using the outturns from the latest data vintage (www.stat.gov.pl accessed in June, 2017), with the last observation available corresponding to 2017:Q1. Obviously, the forecasts for longer horizons are evaluated using fewer outturns (as the longest forecast horizon under consideration is $h = 8$ which corresponds to two-year ahead prediction).

We take into account a number of competing model specifications. Firstly, there are forecasts obtained using the direct approach, i.e. generated from univariate models for Y (the GVA growth rates). Such models amount to assuming $M = 1$ in equations (2)–(4). The specification is denoted by $F(F_i)$ (so $F(1)$ corresponds to a univariate model with one frequency parameter in the Flexible Fourier form in mean of the process; for the univariate models we assume four autoregressive lags ($p = 4$). Another group of models is based on the disaggregation pattern using $\{X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}\}$ variables (hence $M = 4$). The alternative specifications are labelled in a similar way, so $F(2,2,0,0)$ corresponds to a model with $F_1 = 2$, $F_2 = 2$, $F_3 = 2$, $F_4 = 0$. Moreover, $F(0,0,0,0)$ denotes VAR-type deviations from a fixed (time-invariant) mean, corresponding to a BVAR

[2] Details of the prior specification as well as numerical implementation are available from author upon request.

model. For the models we consider different lag structures assuming $p = 2$ and $p = 4$. In particular, we investigate models that allow for the existence of cyclical components for the first two variables only (representing the dynamics of gross value added in industry and trade). For the models based on five variables $\{X_{(1)}, X_{(2)}, X_{(5)}, X_{(6)}, X_{(7)}\}$ we allow for two lags only ($p = 2$), as the number of parameters to be estimated is rather large (compared to the number of available observations). In the multivariate models for disaggregate series, the density forecasts of the aggregate (the growth rates of total GVA) are induced using the procedure outlined above on the basis of the density forecasts for $\{X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}\}$ or $\{X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}\}$.

The forecasts are evaluated using two most popular scoring rules for density forecast evaluation, i.e. the log-predictive score (LPS) and continuous ranked probability score (CRPS); see Gneiting and Raftery (2007). The LPS measure is calculated using natural logs, the values reported are cumulated (not averaged) over the realized forecasts. The CRPS is computed in so-called positive orientation, hence it can be perceived a density generalization of the usual MAE statistics (the values reported are averaged across the realized forecasts). We evaluate the forecasts for each horizon separately, the resulting values of LPS (the higher the better) and CRPS (the lower the better) for selected prediction horizons given in Table 3.

Table 3

Ex-post evaluation of density forecasts for the total GVA dynamics (Y): LPS and CRPS

	LPS					CRPS				
Horizon:	1	2	4	6	8	1	2	4	6	8
$F(0)$ $p = 4$	–11.3	–14.2	–16.2	–14.4	–11.9	0.43	0.63	0.93	0.88	0.70
$F(1)$ $p = 4$	–11.0	–13.7	–15.7	–14.6	–12.4	0.41	0.57	0.87	0.86	0.73
$F(2)$ $p = 4$	–10.8	–13.5	–15.7	–14.9	–13.1	0.39	0.53	0.84	0.87	0.79
$F(0,0,0,0)$ $p = 2$	–11.0	–13.6	–15.5	–14.0	–12.2	0.42	0.58	0.88	0.83	0.73
$F(1,1,1,1)$ $p = 2$	–10.7	–13.5	–15.3	–13.6	–11.8	0.40	0.55	0.82	0.73	0.64
$F(2,2,2,2)$ $p = 2$	–10.4	–13.4	–15.2	–13.6	–11.9	0.38	0.55	0.82	0.75	0.71
$F(2,2,1,1)$ $p = 2$	–10.4	–12.9	–14.9	–13.3	–11.8	0.38	0.52	0.78	0.71	0.70
$F(2,2,0,0)$ $p = 2$	–10.2	–12.5	**–14.5**	**–12.9**	–11.7	0.38	**0.49**	**0.74**	**0.68**	0.68
$F(1,1,0,0)$ $p = 2$	–10.9	–13.3	–15.2	–13.6	–11.8	0.42	0.54	0.81	0.74	0.65
$F(2,0,0,0)$ $p = 2$	**–10.0**	**–12.4**	–14.6	–13.3	–11.7	**0.37**	**0.49**	0.75	0.72	0.66
$F(0,2,0,0)$ $p = 2$	–11.1	–13.7	–15.6	–13.9	–12.3	0.43	0.59	0.91	0.83	0.76

Table 3

	LPS					CRPS				
Horizon:	1	2	4	6	8	1	2	4	6	8
$F(0,0,0,0)$ $p = 4$	–11.0	–13.4	–15.4	–13.8	–11.9	0.42	0.57	0.86	0.80	0.68
$F(1,1,1,1)$ $p = 4$	–10.7	–13.7	–15.6	–13.7	–11.8	0.40	0.57	0.85	0.74	**0.63**
$F(2,2,2,2)$ $p = 4$	–10.6	–13.4	–15.2	–13.4	–11.9	0.39	0.55	0.82	0.71	0.70
$F(2,2,1,1)$ $p = 4$	–10.5	–13.1	–15.0	–13.2	–11.8	0.39	0.54	0.80	0.71	0.70
$F(2,2,0,0)$ $p = 4$	–10.2	–12.5	–14.6	–13.0	–11.6	0.38	0.50	0.75	0.69	0.67
$F(1,1,0,0)$ $p = 4$	–10.8	–13.6	–15.5	–13.6	–11.8	0.41	0.57	0.86	0.75	0.65
$F(2,0,0,0)$ $p = 4$	–10.0	–12.4	–14.7	–13.3	–11.6	0.37	0.50	0.76	0.73	0.65
$F(0,2,0,0)$ $p = 4$	–11.0	–13.6	–15.6	–13.8	–12.0	0.43	0.58	0.89	0.80	0.70
$F(2,2,3,2,1)$ $p = 2$	–11.3	–14.2	–15.4	–13.1	–11.7	0.44	0.64	0.85	**0.68**	0.67
$F(0,0,0,0,0)$ $p = 2$	–11.3	–13.7	–15.4	–13.9	–12.1	0.43	0.58	0.86	0.82	0.71
$F(0,0,0,0,0)$ $p = 4$	–11.2	–13.5	–15.2	–13.7	–11.8	0.43	0.57	0.83	0.78	0.67
$F(1,1,1,1,1)$ $p = 2$	–11.0	–13.2	–14.7	–13.2	**–11.5**	0.43	0.56	0.79	0.73	0.65
$F(2,2,2,2,2)$ $p = 2$	–11.0	–13.8	–15.4	–13.3	–12.0	0.41	0.58	0.84	0.71	0.72
$F(2,2,2,0,0)$ $p = 2$	–11.0	–13.8	–15.5	–13.3	–11.7	0.43	0.61	0.89	0.74	0.69
$F(2,1,1,0,0)$ $p = 2$	–10.9	–13.2	–14.7	–13.4	**–11.5**	0.42	0.56	0.79	0.78	0.66
$F(2,2,0,0,0)$ $p = 2$	–10.8	–13.3	–15.0	–13.4	–11.7	0.42	0.57	0.85	0.77	0.70

Source: own computations using the Central Statistical Office of Poland (GUS) data.

Interestingly, the two criteria under consideration indicate similar models as the best ones. This is not obvious, as it is well known that CRPS and LPS react differently to e.g. tail observations. The results suggest that in all the horizons according to the two criteria, the univariate (direct) forecasts (reported in the first three rows) are never the best ones. There is no clear-cut conclusion as to the best model across all the horizons, however, in a number of cases two specifications dominate: these are $F(2,2,0,0)$, $p = 2$ and $F(2,0,0,0)$, $p = 2$. The models also outperform pure BVARS: F(0,0,0,0) and F(0,0,0,0,0). Moreover, in most cases (except for $h = 8$) the models are better compared to specifications that assume the existence of periodic fluctuations in each variable. The results suggest that it is beneficial to account for the heterogeneity across the sub-aggregates with respect to the existence and properties of the business-cycle type fluctuations. It seems that it is relevant to assume that such regular fluctuations exist for the industry (and perhaps trade), but not necessarily for the other sectors. Moreover, the support for finer disaggregation ($M = 5$) is rather limited. For $M = 4$, in most cases the models with $p = 2$ are preferred to the ones with $p = 4$. We do not attempt to test for 'significance' of the reported differences — as the assumptions underlying such conventional test might be of questionable relevance here. Indeed, the differences are not large in certain cases, but this is not unexpected with short series of quarterly macroeconomic data. We also evaluate performance of point forecasts (predictive expectations), reporting root mean square forecast errors (RMSFE's) in Table 4.

Table 4

RMSFE for point forecasts (predictive expectation) of GVA across models

Horizon	1	2	3	4	5	6	7	8
$F(0)$ $p = 4$	0.74	1.09	1.41	1.64	1.66	1.50	1.24	0.98
$F(1)$ $p = 4$	0.70	0.96	1.25	1.49	1.50	1.37	1.14	0.92
$F(2)$ $p = 4$	0.67	0.87	1.13	1.39	1.40	1.31	1.13	0.96
$F(0,0,0,0)$ $p = 2$	0.73	1.00	1.29	1.57	1.57	1.42	1.21	1.04
$F(1,1,1,1)$ $p = 2$	0.68	0.95	1.24	1.48	1.40	1.15	0.89	0.74
$F(2,2,2,2)$ $p = 2$	0.63	0.93	1.22	1.42	1.37	1.18	1.08	1.04
$F(2,2,1,1)$ $p = 2$	0.65	0.86	1.14	1.35	1.23	1.12	1.02	1.02
$F(2,2,0,0)$ $p = 2$	0.64	**0.81**	1.08	**1.27**	**1.14**	1.06	0.94	0.99
$F(1,1,0,0)$ $p = 2$	0.72	0.93	1.22	1.45	1.38	1.18	0.94	0.81
$F(2,0,0,0)$ $p = 2$	**0.62**	0.82	**1.07**	1.29	1.24	1.15	0.99	0.92
$F(0,2,0,0)$ $p = 2$	0.74	1.02	1.35	1.61	1.59	1.41	1.23	1.12
$F(0,0,0,0)$ $p = 4$	0.72	0.98	1.26	1.52	1.50	1.33	1.09	0.90
$F(1,1,1,1)$ $p = 4$	0.68	0.99	1.32	1.53	1.44	1.18	**0.86**	**0.71**
$F(2,2,2,2)$ $p = 4$	0.66	0.91	1.24	1.41	1.29	1.08	0.96	1.01
$F(2,2,1,1)$ $p = 4$	0.65	0.90	1.20	1.40	1.25	1.10	0.99	1.02
$F(2,2,0,0)$ $p = 4$	0.64	0.83	1.10	1.31	1.21	1.08	0.90	0.93
$F(1,1,0,0)$ $p = 4$	0.71	0.97	1.29	1.53	1.45	1.21	0.94	0.80
$F(2,0,0,0)$ $p = 4$	0.63	0.83	1.09	1.32	1.27	1.17	0.98	0.88
$F(0,2,0,0)$ $p = 4$	0.74	1.00	1.33	1.58	1.53	1.33	1.12	0.99
$F(2,2,3,2,1)$ $p = 2$	0.76	1.11	1.42	1.53	1.30	**1.05**	0.90	0.92
$F(0,0,0,0,0)$ $p = 2$	0.75	1.01	1.27	1.52	1.52	1.37	1.15	0.98

Table 4

Horizon	1	2	3	4	5	6	7	8
$F(0,0,0,0,0)$ $p = 4$	0.73	0.97	1.22	1.45	1.44	1.28	1.04	0.87
$F(1,1,1,1,1)$ $p = 2$	0.74	0.96	1.22	1.42	1.40	1.22	1.00	0.90
$F(2,2,2,2,2)$ $p = 2$	0.71	0.99	1.24	1.48	1.47	1.12	0.96	1.05
$F(2,2,2,0,0)$ $p = 2$	0.73	1.06	1.37	1.61	1.51	1.22	1.05	1.02
$F(2,1,1,0,0)$ $p = 2$	0.72	0.97	1.23	1.42	1.39	1.33	1.08	0.94
$F(2,2,0,0,0)$ $p = 2$	0.72	0.97	1.26	1.51	1.46	1.31	1.09	1.08

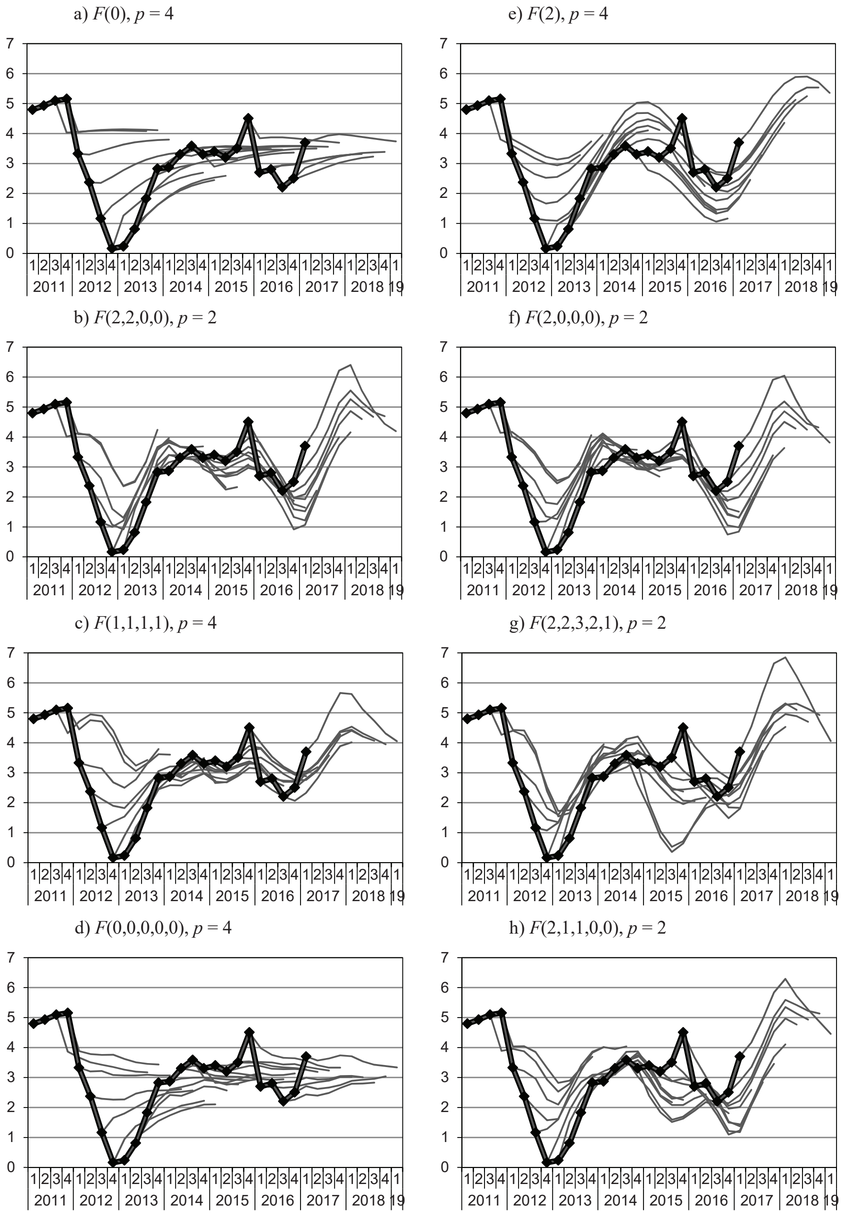
Source: own computations using Central Statistical Office of Poland (GUS) data.

The comparison of point forecast performance leads to very similar conclusions: again, there is no support for the direct use of univariate models.

Important differences across the models can be revealed when analysing whole forecasts paths. Figure 3 depicts the sequences of predictive means, together with the data (as of the most recent vintage). First of all, it is clearly visible that pure autoregressive models (either univariate — case 3a, or multivariate — case 3d) generate forecasts that relatively quickly revert to the time-invariant mean. Such forecast paths are somewhat trivial and it is not obvious whether the results of such kind are capable of providing support for e.g. policy-making.

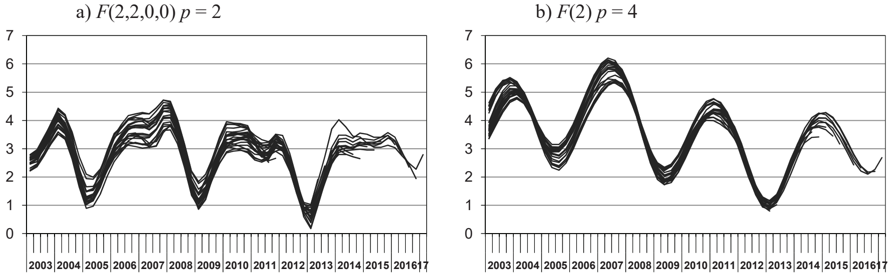
The univariate model with two frequencies in the cyclical component (case 3e) generates non-trivial forecast paths. However, the paths are very regular, directly revealing the properties of trigonometric functions underlying the Flexible Fourier Form — such behaviour might be considered unrealistic from the economic point of view. Interestingly, the models that score well according to the results of Tables 3 and 4, i.e. $F(2,2,0,0)$ $p = 2$ and $F(2,0,0,0)$ $p = 2$ (cases 3b and 3f) generate forecast paths that display variation in the forecast period, but the paths are not as regular as in the univariate model. The results are interesting and might potentially provide guidance for economic decision-making.

In order to provide more thorough comparison of properties of alternative models, we examine a sequence of in-sample estimates of the trends (corresponding to posterior means of $\mu_{t,1}$ for the univariate case and the aggregate of $\mu_{t,i}$'s for the multivariate case); the estimates are depicted in Figure 4. Again, the results provide qualitatively similar conclusions, with one exception: the results from the univariate model (case 4b) are more regular, with peaks and troughs being symmetric in shape. On the other hand, a multivariate model with heterogeneity in sub-aggregates delivers implicit trend estimates that look more realistic (and are characterized by an asymmetric shape of troughs and peaks). As Figure 4 depicts the sequence of estimates taking into account the subsequent revisions, one might conclude that the trend estimates are relatively stable over time.



Source: own computations using the Central Statistical Office of Poland (GUS) data.

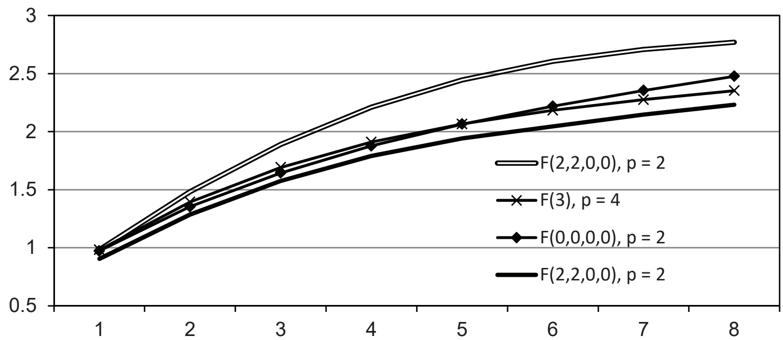
Figure 3. Sequences of point forecast paths (posterior means) for the aggregate quantity Y (GVA)



Source: own computations using the Central Statistical Office of Poland (GUS) data.

Figure 4. Subsequent in-sample estimates of 'cyclical trends' of the aggregate (total GVA growth rates)

One more aspect of density forecast comparison investigated here is that of *ex-ante* forecast uncertainty. It turns out that the average dispersion of the predictive distribution (measured as its standard deviation, which is assumed to exist) from the univariate model with two Fourier components ($F(2)$) is considerably higher compared to other specifications (see Figure 5), especially in longer horizons. The difference is clearly visible when compared to the $F(2,2,0,0)$ $p = 2$ model, which is characterized by lower forecast dispersion (on average).



Source: own computations using the Central Statistical Office of Poland (GUS) data.

Figure 5. Average predictive standard deviation across horizons ($1 < h < 8$) and selected models

Unfortunately, the number of realized forecasts is too small to verify forecast calibration in a reliable way — therefore it is not clear which model is more adequate in this aspect (i.e. whether the *ex-ante* forecast uncertainty matches its *ex-post* counterpart). However, the results depicted in Figure 5 are important. In some cases it might turn out that the disaggregate approach generates satisfactory point forecasts but due to specific stochastic properties of the disaggregate series it might produce e.g. aggregate density forecasts with unacceptably large dispersion — however, this is not an issue for the application under consideration.

The forecasts analyzed here in most cases suggest that the dynamics of the total gross value added for the Polish economy has slowed down reaching a trough in 2016:Q4 or 2017:Q1. Moreover, in most cases the models predict considerably faster growth in subsequent quarters (till mid-2018). On the other hand, quite a high first readout for 2017:Q1 might suggest that the period of growth has occurred one or two quarters earlier, compared to the models' forecasts. This might suggest that the Flexible Fourier pattern is still not flexible enough to capture potential differences or changes in cycle length.

Within the approach used here it is possible to compute contributions of the disaggregate quantities into forecasts of the growth rate of the aggregate quantity — which is appealing for potential users. The possibility is not illustrated explicitly, however, the option is easily available.

In the paper we assume that the disaggregation scheme is given, however, one might think of conducting a random search over a set of possible disaggregation patterns. Such a search could be difficult to conduct in a real-time setup since it requires considerable computational power. However, it is likely that it would be possible to develop an approximate algorithm that would support at least some simple specification search process — e.g. as in Mazur (2015).

Another issue that is left for the further research is the application of the idea of forecast combination. It is likely that it would be possible to obtain even better aggregate forecasts applying some form of forecast pooling. However, the fully formal approach (in the spirit of Bayesian inference pooling) is not possible as one would have to combine implicit forecasts induced from models which differ in terms of the number of endogenous variables.

5. SUMMARY AND CONCLUSIONS

In the paper we investigate the issue of generating density forecasts of macroeconomic quantities using disaggregate data with multivariate models that are capable for accounting for differences in properties of disaggregate series. In particular we make use of models that assume variable-specific type of cyclical fluctuations. The class of multivariate models used here generalizes the approach presented in Lenart and Mazur (2016) into higher dimension.

The reason for consideration of such a question is that some well-established forecasting techniques that make use of disaggregate data have been developed with the purpose of point forecasting in mind. However, nowadays it is obvious that the probabilistic paradigm in forecasting is indispensable. This might come as a challenge, since the probabilistic (or density) prediction leaves much more space for possible model mis-specification (that might be innocuous for the sake of point prediction, but not for the probabilistic prediction). Mimicking the point-forecast oriented models or techniques in density prediction might result in forecasts with unacceptable properties (e.g. forecast dispersion being too small or too large).

Moreover, the issues are unlikely to be solved in a fully formal way: broadly accepted theoretical foundations for model comparison in such a setup have not been developed yet. Consequently, it is necessary to collect empirical evidence that might provide some rough guidance as to the issues considered here. In the paper such an empirical analysis is conducted — the objective is to provide density forecast of the growth rates of total gross value added in the Polish economy (using the data published at quarterly frequency). We develop a predictive experiment which has two important features. Firstly, all the models are re-estimated with each new observation and the estimation uncertainty is accounted for in the density prediction. Secondly, the real-time data setup is followed rather closely: the forecasts are generated using a sequence of twenty two data vintages (constructed based on the archive of quarterly announcements of the Central Statistical Office of Poland).

We consider three groups of competing models: one group represents the direct approach (i.e. the univariate models formulated directly for the quantity of interest). The two remaining groups correspond to 4-variable and 5-variable disaggregation schemes (or sectoral components of the total gross value added).

Results of the predictive experiment indeed indicate that the forecasts based upon the disaggregate approach have better *ex-post* properties compared to the direct forecasts. In particular it seems that the improvements stem from the possibility of assuming different (more regular) fluctuation scheme for the industry (and possibly trade) components, as opposed to other sectors contributing into the total GVA growth in Poland.

Such a conclusion contributes to the on-going debate upon the optimal way of analyzing business cycle fluctuations: it supports the view that the general fluctuation pattern should be implicitly derived based on analysis of the disaggregate data. Moreover, the resulting forecasts have properties that are appealing to practitioners: it is quite easy to make inference not only on the future growth rates, but also on contributions of individual sectors into the total growth. The latter feature is important for those who would like to provide a sort of economic interpretation to the predictive results. Moreover, the results imply rather positive prospects for the growth of the Polish economy, at least for the period 2017:Q2–2018:Q3.

REFERENCES

Clark T.E., Ravazzolo F. (2015), *Macroeconomic Forecasting Performance under Alternative Specifications of Time-Varying Volatility*, Journal of Applied Econometrics, 30(4), 551–575.

Doornik J.A. (2007), *Object-Oriented Matrix Programming Using Ox*, 3rd ed. London: Timberlake Consultants Press and Oxford: www.doornik.com.

Gallant A.R. (1981), *On the Bias in Flexible Functional Forms and an Essentially Unbiased Form: the Fourier Flexible Form,* Journal of Econometrics, 15.

Gneiting T., Raftery A. (2007), *Strictly Proper Scoring Rules, Prediction, and Estimation*, Journal of the American Statistical Association, 102(477).

Harvey A.C., Trimbur T.M., van Dijk H.K. (2007), *Trends and cycles in economic time series: A Bayesian approach*, Journal of Econometrics, 140.

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), 26–67.

Lenart Ł., Leszczyńska-Paczesna A. (2016), *Do market prices improve the accuracy of inflation forecasting in Poland? A disaggregated approach*, Bank & Credit, 47(5), 365–394.

Lenart Ł., Mazur B. (2016), *On Bayesian Inference for Almost Periodic in Mean Autoregressive Models*, Statistical Review/Przegląd Statystyczny, 63(3), 255–271.

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, 212–222.

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), 769–783.

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).

Mazur B. (2015), *Density Forecasts Based on Disaggregate Data: Nowcasting Polish Inflation*, Dynamic Econometric Models, 15, 71–87.

Mazur B. (2017), *Density Forecasts of Polish Industrial Production: a Probabilistic Perspective on Business Cycle Fluctuations*, Institute of Economic Research Working Papers 75/2017.

Stock J.H., Watson M.W. (2014), *Estimating turning points using large data sets*, Journal of Econometrics, 178, 368–381.