EKONOMETRIA ECONOMETRICS 32 • 2011

ISSN 1507-3866

Barbara Batóg
University of Szczecin

FORECASTING OF GROSS DOMESTIC PRODUCT IN POLAND AND ITS REGIONAL COMPONENTS

Summary: The paper presents the forecasts for gross domestic product in Poland and its regional components. The forecasts were set down on the basis of the following models: trends, exponential smoothing and panel data models with an object varying constant. In panel data model there was one exogenous variable – employment in the national economy. The data for employment in a given year are known earlier than data for regional GDP in the same year. Therefore the real values of employment could have been used in calculating the forecasts of GDP. All applied models were well fitted but the forecasts calculated on the base of the panel data model had the smallest relative errors.

Key words: forecasting, regional gross domestic product, trends, panel data model.

1. Introduction

The paper presents the results of forecasting the gross domestic product in Poland and its components, i.e. gross domestic products in voivodeships. In the public statistics there are quite big delays in data. The data on voivodeships' GDP's are available much later than data on Poland's GDP. In the case of gross domestic product that delay could even be equal to two years. Therefore if one tries to forecast the regional GDP there is a gap between a forecasted period and the period of last known data. The aim of the paper is the comparison of the forecast calculated on the basis of a trend model, panel data model and the model of shares of voivodeships' GDP in Poland's GDP. These forecasts are also different because of the hierarchy of them – in the latter the forecast for Poland is superior and in the former is not.

2. The model with intercept varying over objects

Forecasting on the basis of trends is described in many books, for example Zeliaś [1997], Dittmann [2004]. In the case of forecasting the shares of voivodeships' GDP in Poland's GDP, the trends as well as exponential smoothing models were used (see [Gardner 1985]). The model where the intercept varies over objects (voivodeships) was applied to panel data (see [Judge et al. 1985; Cameron, Trivedi 2005; Maddala 2006].

The model where the intercept varies over objects (for ith object) could be written as:

$$y_i = \beta_{1i} j_T + x_{ki}\beta_k + e_i,$$

where: $y_i = [y_{i1}\ y_{i2} ... y_{iT}]'$,
$e_i = [e_{i1}\ e_{i2} ... e_{iT}]'$ – random errors,
$j_T = [1\ 1 ... 1]'$, $(T\text{x}1)$,
x_{ki} – matrix of values of the explanatory variables except for the constant $(Tx(K-1))$,
T – number of periods,
N – number of objects,
K – number of explanatory variables,
β_k – slope coefficients for explanatory variables except for the constant,
β_{1i} – object varying intercepts.

It is assumed that random errors e_i (i = 1, 2, ..., N) for every object and period have expected values equal zero, constant variance and are not correlated with different objects. If the number of objects is not too large, this model could be estimated for all objects jointly. The estimates of intercept that varies over objects could be obtained by introducing dummy variables into the model. The ith dummy variable equals one for ith object (voivodeship) and zero for others. With the above assumptions the ordinary least squares estimator is the best linear estimator.

3. Empirical data on GDP in Poland and GDP in Polish voivodeships

The data on Poland's gross domestic product in 1995-2009, voivodeships' gross domestic product in 1995-2007, employment in the national economy (main place of work) in 2002-2008[1] came from the Local Data Bank of the Central Statistical Office.

Figure 1 presents the tendency of Poland's gross domestic product in 1995-2009. In this whole period the GDP was increasing. Only in 2001-2003 was the increase slower than in 1995-2000 and 2004-2009.

Figure 2 presents the gross domestic product in every voivodeship in 1995-2007. This period is shorter than the period for Poland presented in Figure 1 because regional data has a two years delay. The differentiation of GDP is much bigger in 2007 than in 1995. The difference in rates is visible. The biggest growth rates were noted in Mazowieckie and Śląskie voivodeships and the lowest ones in Opolskie, Warmińsko-mazurskie, Świętokrzyskie and Podlaskie voivodeships.

[1] Data on employment before 2002 was not comparable with data after 2002.

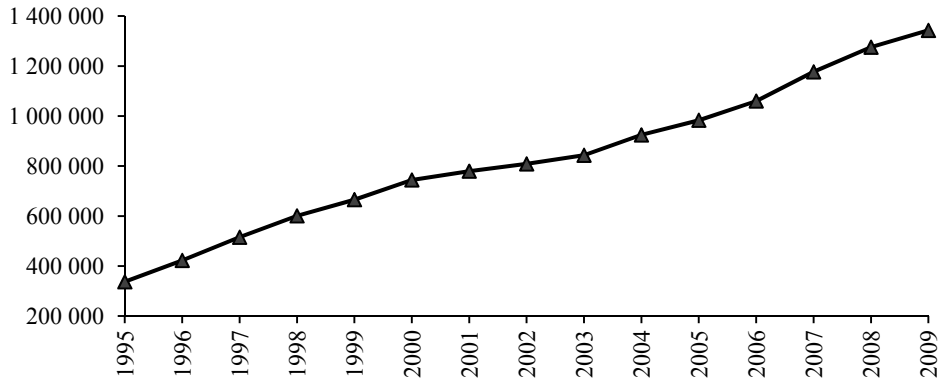

Figure 1. Gross domestic product in Poland in 1995-2009 (millions PLN)

Source: Local Data Bank of Central Statistical Office.

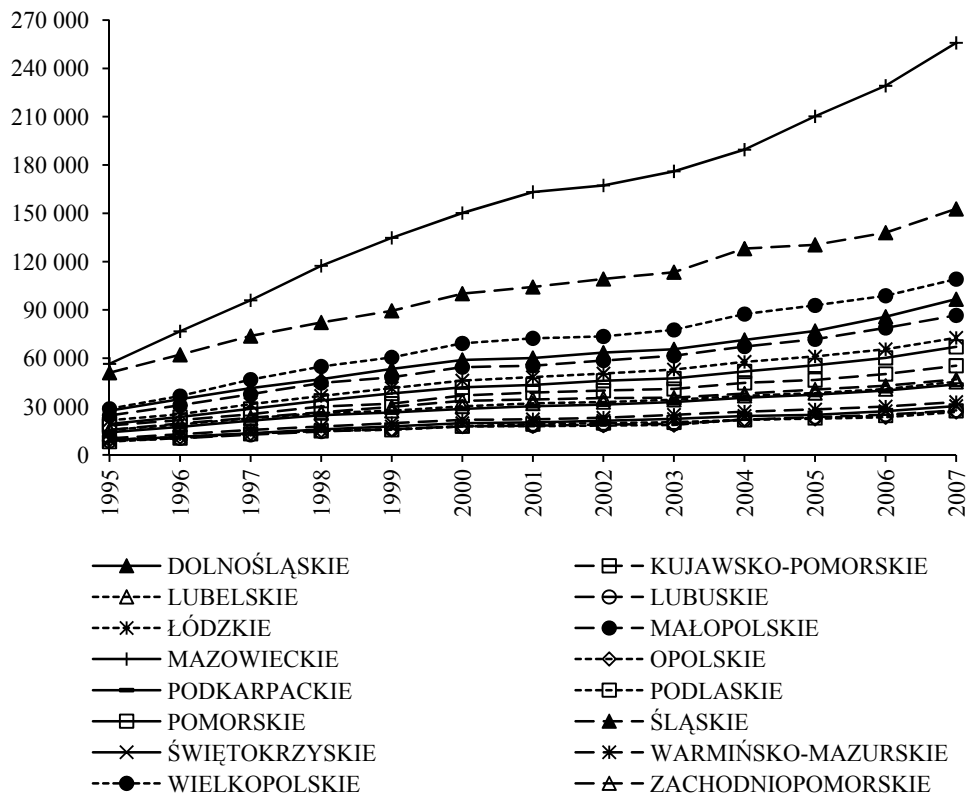

Figure 2. Gross domestic product in Polish voivodeships in 1995-2007 (millions PLN)

Source: Local Data Bank of Central Statistical Office.

4. Forecasts of total and regional GDP

First the trend models for total GDP and for regional GDP were estimated. On the base of the estimated models the forecasts for 2008 were computed. Table 1 presents relative prediction errors (*ex ante*) for these forecasts. The best analytic form for all voivodeships was a linear trend. The parameters of linear trends were statistically significant and the coefficients of determination (R^2) were very high (about 90%) for all voivodeships. The relative prediction errors also had very good properties – their values were below 5%.

Table 1. Relative prediction errors (*ex ante*) of GDP forecasts based on trend models for voivodeships for 2008

Voivodeship	Relative prediction errors (*ex ante*) [%]
Dolnośląskie	4.6
Kujawsko-pomorskie	3.6
Lubelskie	3.7
Lubuskie	3.9
Łódzkie	3.3
Małopolskie	3.8
Mazowieckie	4.1
Opolskie	4.6
Podkarpackie	3.6
Podlaskie	4.1
Pomorskie	3.8
Śląskie	2.9
Świętokrzyskie	3.9
Warmińsko-mazurskie	3.4
Wielkopolskie	3.9
Zachodniopomorskie	5.0

Source: own calculations based on the data from Local Data Bank of Central Statistical Office.

The forecasts of regional GDP for 2008 were summed up and compared to the real value of the total GDP in 2008. The relative error (*ex post*) of these summed up forecasts was equal to 5.72%.

In the case of the estimation of panel data model (with object varying intercept) employment (main place of work) was the explanatory variable. The tendencies of employment (main place of work) in voivodeships in 2002-2008 are presented in Figure 3.

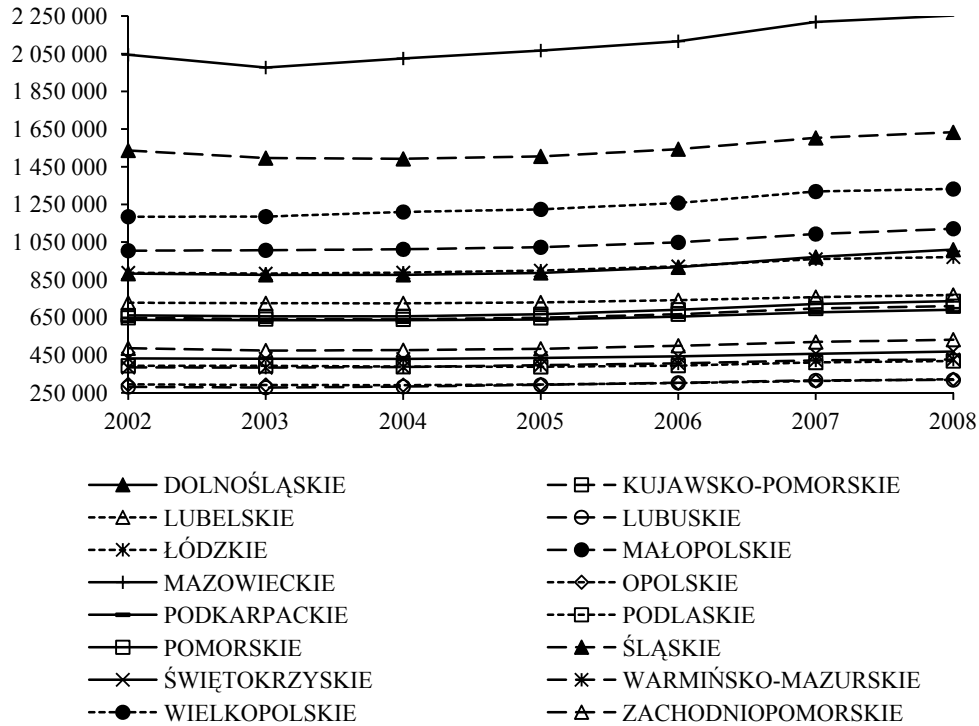

Figure 3. Employment in the national economy (in main place of work) in voivodeships in 2002-2008 (persons)

Source: Local Data Bank of Central Statistical Office.

The estimated parameters of panel data model (with object varying intercept) with one explanatory variable (employment) are presented in Table 2. The object varying intercepts were represented by dummy variables z_i. The number of dummy variables should be equal to 16, i.e. the number of voivodeships. However, the estimation is possible only when the number of dummy variables is reduced by 1 and the model is reparameterized and incorporates the constant and does not incorporate one of the dummy variables – here z_{16} which represents Zachodniopomorskie voivodeship. The estimated model was characterized by very high coefficient of determination $R^2 = 98.9\%$. Only one of the dummy variables was not statistically significant. It was the dummy that represented the Świętokrzyskie voivodeship. The statistical insignificance of the dummy variable means that the parameter for Świętokrzyskie voivodeship is practically equal to the constant and also that intercept for Świętokrzyskie voivodeship is equal to the intercept for the Zachodniopomorskie voivodeship.

On the base of the estimated model (presented in Table 2) the forecasts of regional GDP for 2008 were calculated. The real values of employment in the national economy (main place of work) for 2008 were used. It was done this way because, as

Table 2. The estimated parameters of panel data model (with object varying intercept) $R^2 = 0.989$

	Estimated parameters	Standard error	p level
Constant	–110 974.5	9 280.14	0.000
Employment	0.308	0.02	0.000
z1	–89 805.8	8 188.72	0.000
z2	–44 989.7	4 393.51	0.000
z3	–77 049.3	5 484.35	0.000
z4	43 761.9	4 811.02	0.000
z5	–107 814.9	8 270.01	0.000
z6	–135 615.7	10 432.21	0.000
z7	–322 788.4	29 300.66	0.000
z8	41 210.9	4 732.56	0.000
z9	–51 316.2	4 261.22	0.000
z10	12 225.3	3 598.33	0.001
z11	–42 180.9	4 634.60	0.000
z12	–231 041.4	19 363.25	0.000
z13	1 148.9	3 285.24	0.727
z14	16 515.8	3 581.44	0.000
z15	–177 683.6	13 962.01	0.000

Source: own calculations on the basis of data from Local Data Bank of Central Statistical Office.

Table 3. The relative prediction errors (*ex ante*) of forecasts calculated on the basis of panel data model (with object varying intercept)

Voivodeship	Relative prediction errors (*ex ante*) [%]
Dolnośląskie	5.6
Kujawsko-pomorskie	9.5
Lubelskie	12.1
Lubuskie	19.2
Łódzkie	7.5
Małopolskie	6.2
Mazowieckie	2.6
Opolskie	20.2
Podkarpackie	11.8
Podlaskie	19.6
Pomorskie	8.1
Śląskie	3.8
Świętokrzyskie	17.1
Warmińsko-mazurskie	16.0
Wielkopolskie	5.1
Zachodniopomorskie	7.2

Source: own calculations on the basis of data from Local Data Bank of Central Statistical Office.

it is mentioned at the beginning of the paper, data on employment is available from public statistics earlier than data on regional GDP. The relative prediction errors (ex ante) of forecasts calculated on the basis of panel data model (with object varying intercept) are presented in Table 3.

For several voivodeships the errors from Table 3 are higher than 15%. Nevertheless when these regional forecasts were summed up and compared to the real value of Poland's GDP it turned out that the relative error (*ex post*) of Polish GDP forecast was equal to only 0.46%.

For the purpose of comparison the second panel data model (with object varying intercept) was estimated (the time variable was the explanatory variable). The relative errors (*ex post*) of forecasts for regional GDP calculated on the base of this model were higher than in the case of forecasts for regional GDP calculated on the base of individual (separate for each voivodeship) trend models. The relative error (*ex post*) of forecasts for the total GDP (as a sum of regional GDP forecasts) was equal to 5.65%.

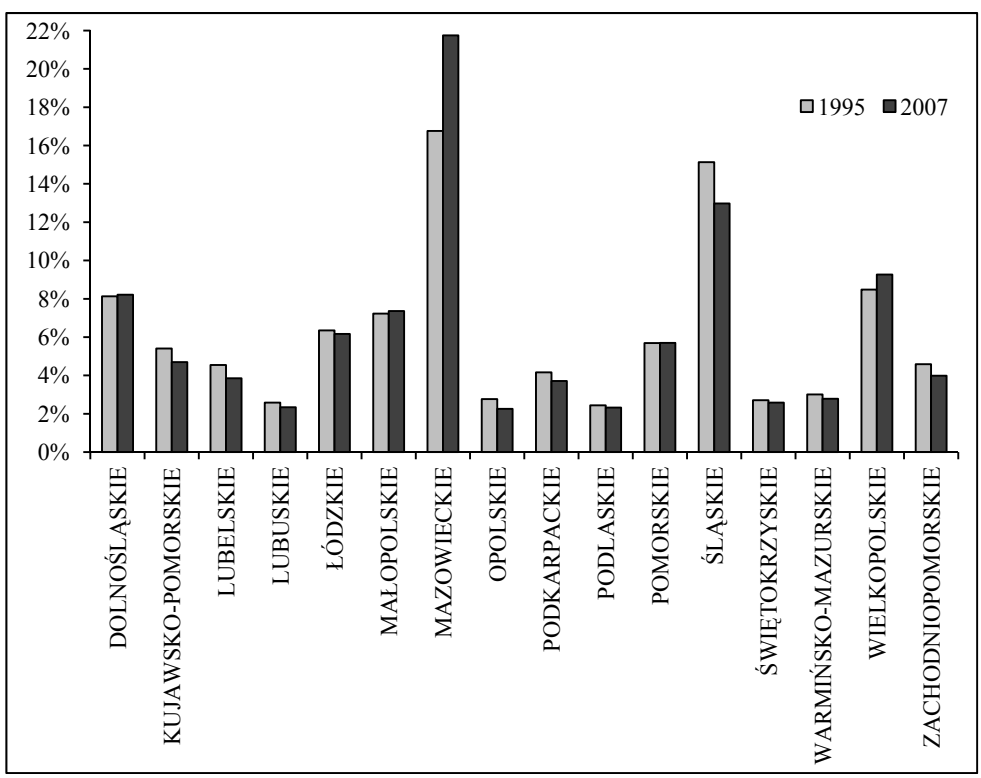

Figure 4. Shares of voivodeships' GDP in Poland's GDP [%]

Source: own calculations on the basis of data from Local Data Bank of Central Statistical Office.

Then the inverse approach was applied. This time Polish GDP was superior to the voivodeships' GDP. Therefore the shares of regional GDP's in total GDP were calculated. The shares for 1995 and 2007 are presented in Figure 4. In most voivodeships these shares have changed only a little. The situation is different in the case of the Mazowieckie voivodeship – the increase is visible – and in the case of the Śląskie voivodeship – the decrease is visible.

The forecasts of shares of regional GDP in total GDP were calculated by means of two methods. For eleven voivodeships these shares were characterized by strong linear regularities, so the linear trend models were estimated and then the forecasts of shares were computed. For the remaining five voivodeships the Holt linear exponential smoothing model was applied to calculate the forecasts. In order to obtain the forecasts of regional GDP for 2008, the forecasts of shares of regional GDP in total GDP were multiplied by the forecast for Polish GDP calculated on the base of linear trend. This multiplication is connected with the idea of internally consistent forecasts where the sum of forecasts of the regional GDP has to be equal to the forecast of Polish GDP [Guzik 1976].

Table 4. The forecasts of shares of regional GDP in total GDP and obtained on this basis the forecasts of regional GDP for 2008

Voivodeship	Shares' forecasts	GDP's forecasts
Dolnośląskie	0.083*	100041.8
Kujawsko-pomorskie	0.046	55725.9
Lubelskie	0.037	44780.4
Lubuskie	0.023*	27968.4
Łódzkie	0.062*	74003.8
Małopolskie	0.074*	88293.0
Mazowieckie	0.225	270334.7
Opolskie	0.021	25343.1
Podkarpackie	0.037	44005.8
Podlaskie	0.023	27674.3
Pomorskie	0.057*	68033.3
Śląskie	0.127	152990.6
Świętokrzyskie	0.025	30607.5
Warmińsko-mazurskie	0.028	33529.3
Wielkopolskie	0.095	114604.4
Zachodniopomorskie	0.040	47674.3

* – forecasts calculated on the basis of Holt exponential smoothing model

Source: own calculations on the basis of data from Local Data Bank of Central Statistical Office.

The relative error (*ex post*) for the forecast for Polish GDP calculated on the base of linear trend was equal to 5.69%. The shares' forecasts and the GDP forecasts are presented in Table 4.

5. Conclusion

It can be stated that the smallest errors of forecasts for regional GDP were obtained in the case of the panel data model (with object varying intercept). The possibility of using the real values of the explanatory variable – employment in the national economy (in the main place of work) – had a positive influence on the quality of these forecasts. However, it was a special situation not always existing in practice. In the case of the lack of real values of the explanatory variable the other two methods can be applied. The forecasts calculated by means of the two remaining methods had errors only a little bit higher. Therefore they are also very useful. Because of the good quality of forecasts computed on the base of the used method there is also the possibility of combining these forecasts. However, it should be remembered that the gross domestic product is the additive variable, so there is a strict relationship (sum) between the values on different levels of aggregation. The presented forecasts fulfill this assumption.

The forecasts of regional GDP can be used by local governments. They are able to make better decisions to the benefit of the local community on the basis of better forecasts.

Literature

Cameron A.C., Trivedi P.K., *Microeconometrics. Methods and Applications*, Cambridge University Press, Cambridge 2005.

Dittmann P., *Prognozowanie w przedsiębiorstwie*, Oficyna Ekonomiczna, Kraków 2004.

Gardner E.S., *Exponential smoothing. The state of the art*, "Journal of Forecasting" 1985, Vol. 4, Issue 1.

Guzik B., *O wewnętrznej zgodności prognoz ekonometrycznych. Sformułowania problemu*, „Przegląd Statystyczny" 1976, nr 3.

Judge G.G., Griffiths W.E., Hill R.C., Lütkepohl H., Lee T-C., *Theory and Practice of Econometrics,* John Wiley and Sons, 1985.

Maddala G.S., *Ekonometria*, Wydawnictwo Naukowe PWN, Warszawa 2006.

Zeliaś A., *Teoria prognozy*, PWE, Warszawa 1997.

PROGNOZOWANIE ZMIENNEJ AGREGATOWEJ ORAZ JEJ SKŁADOWYCH NA PRZYKŁADZIE PRODUKTU KRAJOWEGO BRUTTO

Streszczenie: W artykule przedstawiono wyniki prognozowania produktu krajowego brutto ogółem oraz w ujęciu regionalnym. Wykorzystano model trendu, modele wyrównywania wykładniczego oraz modele panelowe ze zmieniającym się po obiektach wyrazem wolnym. W przypadku modeli panelowych zmienną objaśniającą była liczba pracujących w gospodarce narodowej. W statystyce publicznej liczba pracujących w poszczególnych województwach w danym roku jest znana znacznie wcześniej niż regionalny produkt krajowy brutto dla tego samego roku. Dlatego też przy wyznaczaniu prognoz PKB wykorzystano rzeczywiste wartości liczby pracujących. Wszystkie wykorzystane modele charakteryzowały się wysoką jakością dopasowania, natomiast najmniejszymi błędami były obarczone prognozy wyznaczone na podstawie modelu dla danych panelowych.