ANALYSIS OF SEASONALITY IN TIME SERIES OF LOANS IN AGRICULTURE

Ondřej Šimpach

Abstract

The aim of the paper is to analyse whether there is a seasonality in loans that are provided to agricultural, forestry and fishery enterprises and to project the demand for loans up to the end of the year 2023. We use monthly time series of loans to agriculture for last 20 years – from 01/2002 to 01/2023 (253 observations). We apply two methods for time series analysis and compare them based on RMSE and according how realistic are their forecasts.

First method, deterministic decomposition on trend and seasonal part, revealed that there was statistically significant trend and seasonality in February, March and April. Also stochastic method, SARIMA(1,1,0)(1,0,1)$_{12}$ had statistically significant seasonal parameters. Tests revealed statistically significant seasonality and moving seasonality. Despite that deterministic method has lower RMSE than retropolated SARIMA, its forecast seems to be overestimated. When SARIMA without retropolation is applied, the forecast is the most realistic. We can expect loans to agriculture volume to grow up to 66 milliard CZK at the end of 2023.

Key words: decomposition, loans, seasonality, time series

JEL Code: C22, Q14

Introduction

Commercial loans are one of the sources of financing of agricultural producers. Agricultural sector used to be evaluated as risky by the banks due to the character of agricultural production that depends on biological processes. Besides "agriculture related small and micro enterprises are not easy to get loans from commercial banks because they don't have collaterals" (Yuan and Yang, 2018).

Loan availability for agricultural companies improved after the entrance of the Czech Republic to the EU because of the provision of subsidies from European Agricultural Fund for Rural Development and provision of direct payments on agricultural land. Commercial banks started to make available bridging loans for farmers for pre-financing of the investments. Co-financing of investments was then seen by banks as less risky when the subsidies were provided after the realization of the investment. According to the findings of Toth, Rabek and Strapekova (2020) based on the case study in Slovakia, the banks consider the CAP subsidies to be a stable income factor and good collateral for loans.

Besides investment loans, that are usually provided for long-term or middle-term, short-term loans for operations are also sometimes necessary. Seasonality is a typical feature of agriculture. The pace of works changes throughout the year as same as the financial cash flow. While costs are incurred mainly in the spring during pre-sowing preparation and sowing, the proceeds from the sold production are not available until late summer or early autumn. This may result in the necessity of short-term loans during the year.

Košovská and Váryová (2017) examined the loans in the Slovak Republic in 2008–2015. They found out that short-term loans prevailed and that they were used mainly for operations of agricultural companies. Same conclusion was done by Toth, Rabek and Strapekova, 2020).

Fecke, Feil and Musshoff (2016) examined the determinants of the loan demand in agriculture. They found out that amongst others, the interest rate, gross value added in agriculture, grace period and farmers' business expectations have significant effects on the loan demand in agriculture. The interest rate had significant negative effect; the others had significant positive effect.

Our focus is on the total amount of loans to agricultural and related sectors.

1 Data and Methods

The aim of the paper is to analyse whether there is a seasonality in loans that are provided to agricultural, forestry and fishery enterprises and to project the demand for loans for one year. The data about clients' loans to agriculture, forestry and fishery were downloaded from ARAD time series database of Czech National Bank (CNB, 2023a). We used monthly time series of loans to agriculture for last 20 years (from 01/2002 to 01/2023), so we had 253 observations.

We applied deterministic decomposition method that decompose the time series on trend and seasonal part using time vector and 11 vectors of dummy variables (taking value 1 in particular month and 0 in other cases). There could not be implemented all 12 dummy variables for 12 months because of perfect multicolinearity among them. Therefore, linear regression model was estimated by Least Square Method in the following form (1).

$$y_t^* = \beta_0^* + \beta_1^* t + \gamma_1^* D_{1t} + \gamma_2^* D_{2t} ... + \gamma_{11}^* D_{11t} + \varepsilon_t \,, \qquad (1)$$

where y_t is the explained variable – loans to agriculture, forestry and fishery in year t, t is the time vector (taking values from 1 to 253), D_{1t} to D_{11t} are dummy variables for January to November, β_0 is constant and β_1, γ_1 to γ_{11} are parameters. ε is the error term. Asterisks marks that the parameters are of auxiliary model and have to be then recalculated to original form of the equation that contains 12 months (2).

$$y_t = \beta_0 + \beta_1 t + \gamma_1 D_{1t} + \gamma_2 D_{2t} + ... + \gamma_{12} D_{12t} + \varepsilon_t \qquad (2)$$

The recalculation is done using auxiliary variable $\hat{s}$ that is calculated as (3).

$$\hat{s} = \frac{\hat{\gamma}_1^* + \hat{\gamma}_2^* + ... + \hat{\gamma}_{11}^*}{12} \qquad (3)$$

The constant is recalculated as $\hat{\beta}_0 = \hat{\beta}_0^* + \hat{s}$, parameter for trend is the same ($\hat{\beta}_1 = \hat{\beta}_1^*$). Parameters for dummy variables are $\hat{\gamma}_j = \hat{\gamma}_j^* - \hat{s}$, where j = 1, 2, ... 11 and $\hat{\gamma}_{12} = -\hat{s}$. If the parameters for particular months are statistically significant, we can conclude that seasonality is an important feature of the time series of loans to agriculture. Estimated equation in MS Excel is then used for projection for next months until the end of the year 2023.

Another way, how to examine and model seasonality in the time series is using Box and Jenkins (1970) methodology. Particularly we applied SARIMA (Seasonal Autoregressive Moving Average) in SW EViews 10. ARIMA(p, d, q) models contain autoregressive (AR) and moving average (MA) processes, where p is the order of AR term, d is the number of non-seasonal differences and q is the order of MA term. The order p of AR process and order q of MA process are determined by Autocorrelation function (ACF) and Partial Autocorrelation function (PACF) that are plotted in correlograms. When there is a seasonal component, a model is in a form of SARIMA(p, d, q)(P, D, Q)$_S$, where P is the order of SAR model, D is order of seasonally differencing, Q is the order of SMA model, and S means the periodicity. Model is then written as (4).

$$y_t = \sum_{i=1}^{p} \varphi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{i=1}^{P} \Phi_i y_{t-S} + \sum_{j=1}^{Q} \Theta_j \varepsilon_{t-S} \,, \qquad (4)$$

where φ and Φ are parameters for lagged explained variable y (i.e. it is explanatory variable). θ and Θ are parameters for lagged error term ε. Estimated optimal model is then used for projection of loans volume until the end of the year 2023.

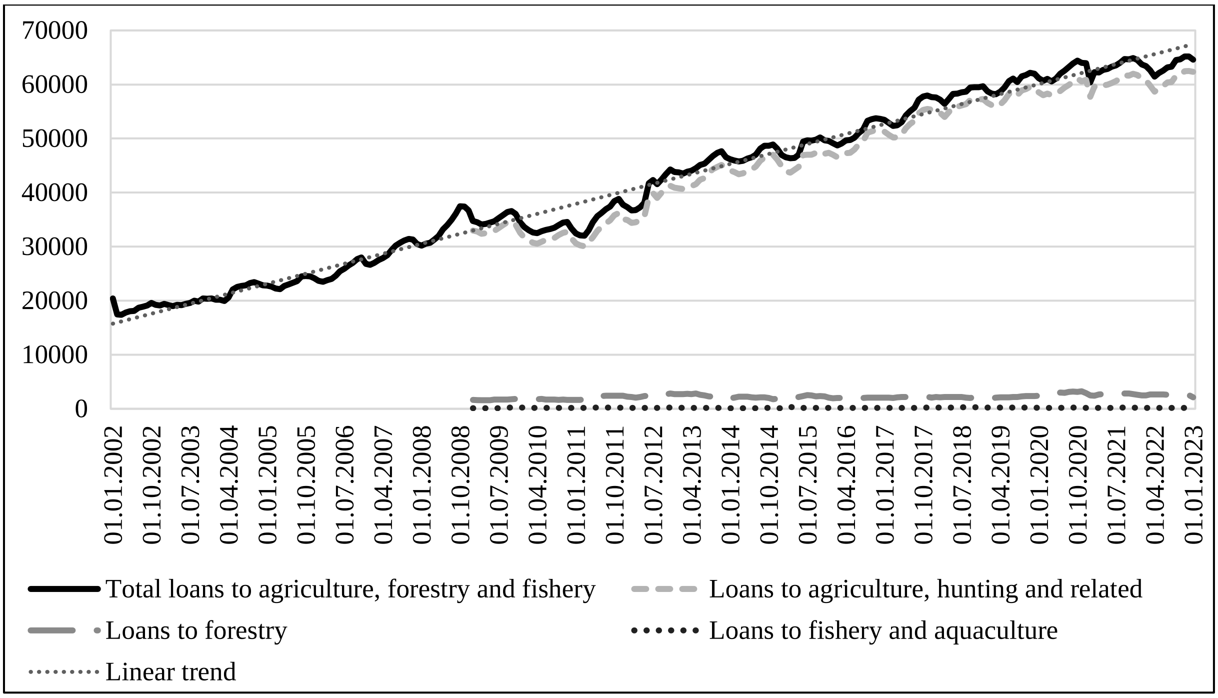
The results of both approaches are compared based on Root Mean Squared Error (RMSE) and their forecasts are discussed.

2 Results

First, the development of the volume of loans was plotted. As can be seen from Fig. 1, the amount of loans to agriculture, forestry and fishery increased continuously during the observed period 2002–2023. There is a clear increasing trend with certain seasonal deflections which we test whether they are significant. The volume of loans increased more than 3 times since the beginning of the period.

Majority of loans was provided to agriculture, hunting and related activities. The loans to forestry and to fishery and aquaculture were negligible in comparison to them. Nevertheless, this amount is not significant in comparison to loans to other sectors. Loans to agriculture, forestry and fishery contributes to the total amount of loans in economy only from 1.5 to 2.3%.

Fig. 1: Loans to agriculture, forestry and fishery



Source: CNB (2023a), own elaboration

Regarding the interest rates, they are observed by the CNB only for the division on financial and non-financial companies, not according to the sectors. CNB (2023b) Therefore, we present the interest rates for non-financial companies on 31st December of particular year. There were around 4% from the beginning between 2004–2006. They outreached 5.0% in 2006 and started to slowly decrease on 1.74% in 2015 which was the minimum. The interest rates increased since that, up to 3.62% in 2019 and dropped to 1.86% one year later. Due to Covid-19 pandemic, there was sharp increase to 4.45% in 2021 and 8.97 % in 2022. Hence, the situation with loans availability for agricultural companies worsen. Surprisingly, however, there was no decrease in the volume of loans provided.

The results for deterministic decomposition of the time series of loans on trend and seasonal part are displayed in Tab 1. Coefficient of determination and adjusted coefficient of determination were high: R^2 = 98.6%, adjusted R^2 = 98.5%. This means that the changes in loans volume are explained from almost 99.0% by the trend and season which is a good result. P-value of the F test was almost zero, so the model as a whole was statistically significant.

The results show that the constant and linear trend are statistically significant. Autonomous demand for loans is in height of 15 584 mil. CZK. The volume of loans increases every month on average by 204.8 million CZK.

When the estimated coefficients are with negative sign, it means that the amount of loans is lower than in December. If the sigh is positive, then on average a higher amount of loans is provided in this month than in December. The statistically significant seasonality was found in February, March and April. The amount of loans is significantly lower than in December by more than 1000 mil. CZK. Other months are statistically insignificant. The estimated coefficients were recalculated using auxiliary variable $\hat{s}$ = -272.38. At the beginning of the year (spring) there is significantly lower amount of loans provided to agriculture.

Tab. 1: Deterministic decomposition of time series 2002–2023

Variable	*Estimated coefficients*	*Error*	*t-value*	*p-value*		*95% confidence interval*		*Recalculated coefficients*
Constant	15856.30	452.39	35.05	*0.000*	***	14965.137	16747.459	15583.91
t	204.84	1.58	129.32	*0.000*	***	201.717	207.957	204.84
D_{1t}	-641.11	560.92	-1.14	0.254		-1746.074	463.846	-368.73
D_{2t}	-1026.32	567.57	-1.81	*0.072*	*	-2144.377	91.740	-753.93
D_{3t}	-1346.12	567.53	-2.37	*0.018*	**	-2464.093	-228.141	-1073.73
D_{4t}	-1294.19	567.49	-2.28	*0.023*	**	-2412.089	-176.285	-1021.80
D_{5t}	-927.01	567.46	-1.63	0.104		-2044.846	190.827	-654.63
D_{6t}	-308.05	567.43	-0.54	0.588		-1425.831	809.729	-35.67
D_{7t}	-32.10	567.41	-0.06	0.955		-1149.834	1085.630	240.28
D_{8t}	121.91	567.39	0.21	0.830		-995.779	1239.606	394.30
D_{9t}	630.83	567.37	1.11	0.267		-486.828	1748.496	903.22
D_{10t}	847.48	567.36	1.50	0.137		-270.157	1965.123	1119.87
D_{11t}	706.061	567.353	1.244	0.215		-411.567	1823.688	978.44
D_{12t}								272.38

Source: own elaboration based on data from CNB (2023a)

First, the time series was examined by F-tests for seasonality. It was revealed that there is seasonality present at the 0.01 level of significance (F-value = 69.425). Moving seasonality test found out that there is also moving seasonality present at 0.1 level of significance (F-value = 5.686).

After the calculation of ACF and PACF functions an optimal model was estimated in a form of SARIMA(1,1,0)(1,0,1)$_{12}$, where variable IMP (unit impulse) is equal to 1 on positions 2009M01 and 2021M01; and IMP is equal to -1 on positions 2012M06, 2015M06, 2021M02.

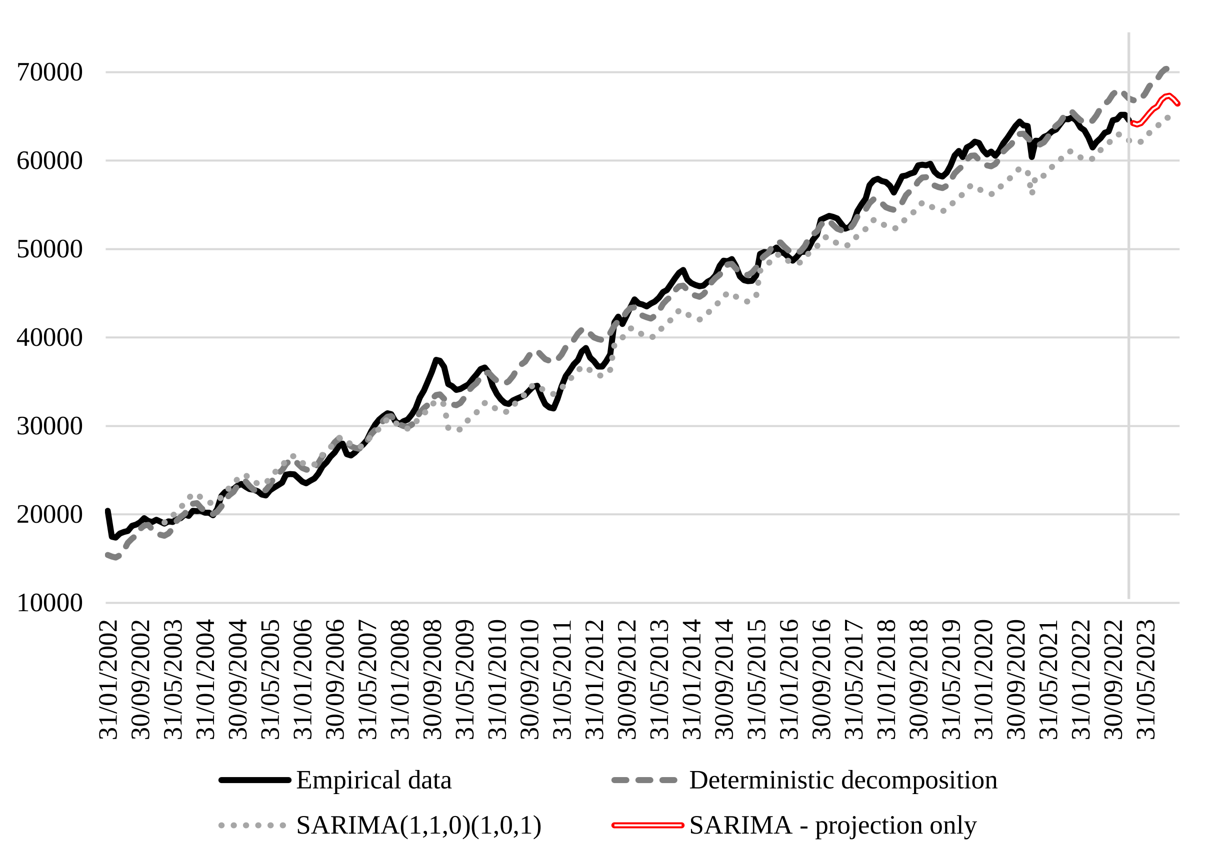
When difference between empirical data and fitted data are too high at some observations, those shocks cause non-normality of the residual component. Therefore, the unit exogenous shocks (IMP) are used at those observations, and the model apply empirical data instead of fitted here. This can ensure the normality of residual component. In our case, application of the impulses helped to create normal distribution of residuals while not causing autocorrelation. The JB statistics was 3.559 with p-value 0.169 which lead to the non-rejection of null hypothesis (H_0: data is normally distributed).

Breusch-Godfrey serial correlation test had F-statistic equal to 0.133 with p-value 0.876, so the null hypothesis was not rejected (H_0: there is no autocorrelation of residuals). ARCH test for heteroscedasticity had F-statistic equal to 0.230 with p-value 0.632, which lead to the non-rejection of H_0: there is homoscedasticity – constant and finite variance of residuals.

3 Discussion

Both used methods found seasonality, the deterministic method identified only certain months (February, March and April). Results of both models (fitted data versus empirical data) are displayed at Fig. 2. It can be seen that up to year 2008, both models fit the empirical data well. Then the economic crisis arrived and the difference between the models and between fitted and empirical data increased. Deterministic decomposition approached to real data again in 2012. However, the consolidation lasted only to year 2017. The fitted data is overestimated since the end of 2021 which caused the overestimation of the amount of loans until the end of the year 2023. On the other hand, SARIMA model underestimated the real data and also the projection. The probable development, however, would be in the middle. If we do not retropolate the data in SARIMA model, but only project next 12 months, then the projection is more realistic. Root mean squared error (RMSE) of SARIMA model was higher (2954.102) than of deterministic model (1790.571), which could lead us to the conclusion that SARIMA is worse model, but the projection without retropolated data is close to reality.

Fig. 2: Projection of loans to agriculture, forestry and fishery – comparison of models



Source: own elaboration based on data from CNB (2023)

Box-Jenkins methodology is often used to project prices – for example Pechrová and Šimpach (2017) used ARIMA model for modelling consumer eggs prices – optimal model based on the diagnosis by ACF and PACF functions was chosen ARIMA(1,0,0). Loans are projected by SARIMA model for example by Chen (2011) who found out that loans of college library have seasonal variation characteristics for which is this model suitable. "Two peaks are shown in the beginning and the end of the semesters, as the two valleys are shown in the two holiday months." (Chen, 2011). Sabu and Kumar (2020) used SARIMA and neural network model to project the prices of arecanuts and found out that based on RMSE, the SARIMA model performed worse. Navarro-Esbrí, Diamadopoulos and Ginestar (2002) compared SARIMA model with a prediction technique based on non-linear dynamics to forecast municipal solid waste management. They found out that non-linear forecasting technique gave comparable results to the SARIMA methodology. The best projection was given by SARIMA model in our case. Therefore, in the future research we will try more methods for time series projection and compare them based on RMSE.

Conclusion

The aim of the paper was to analyse whether there is a seasonality in loans that are provided to agricultural, forestry and fishery enterprises and to project the demand for loans up to the end of the year 2023. We applied two methods and compared them based on RMSE and according how realistic are their forecasts. First method was deterministic decomposition on trend and seasonal part and second stochastic SARIMA. Deterministic method found statistically significant seasonality in February, March and April. Also stochastic method, SARIMA(1,1,0)(1,0,1)12 had statistically significant seasonal parameters (SAR, SMA). Tests revealed statistically significant seasonality and moving seasonality in the time series of loans to agriculture, forestry and fishery. Despite that the deterministic method had lower RMSE than retropolated SARIMA, its forecast seemed to be overestimated. When SARIMA without retropolation was applied, the forecast was the most realistic.

Both models projected increase of the loans volume that is in line with the long-term trend of the development of the loans to agriculture. However, deterministic decomposition probably overestimated possible development because it projected increase of the loans volume up to 70 000 mil. CZK at the end of the year 2013. Contrary to that SARIMA model forecasted that the volume will be almost 67 000 mil. CZK in December 2013, that is more realistic and comparable to the current real development.

Challenge for future research is to examine the loans according to their length as we can suppose that short-term loans are more seasonal than long-term or mid-term loans. Also, more methods of time series projection can be compared.

Acknowledgment

The paper was supported by long-term institutional support of research activities by the Faculty of Informatics and Statistics of Prague University of Economics and Business.

References

Box, G. & Jenkins, G. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Chen, L. H. (2011). SARIMA Model Used to Forecast the Loan Amounts of College Library. *Proceedings of Recent Advance in Statistics Application and Related Areas*, 679–682.

Czech National Bank (CNB). (2023). *Klientské úvěry podle odvětví (CZ-NACE) - podle sekcí (Kč+cizí měna).* Databáze časových řad ARAD. https://www.cnb.cz/cnb/STAT.ARADY_PKG.PARAMETRY_SESTAVY?p_sestuid=51031&p_strid=AABBAC&p_lang=CS

Czech National Bank (CNB). (2023). *Tabulka B1.1.3: Úrokové sazby korunových úvěrů poskytnutých bankami nefinančním podnikům v ČR – nové obchody (%).* Databáze časových řad ARAD. https://www.cnb.cz/cnb/STAT.ARADY_PKG.PARAMETRY_SESTAVY?p_sestuid=58843&p_strid=AAABAA&p_lang=CS

Fecke, W., Feil, J.-H. & Musshoff, O. (2016). Determinants of loan demand in agriculture: Empirical evidence from Germany. *Agricultural Finance Review*, *76*(4), 462–476. https://doi.org/10.1108/afr-05-2016-0042

Košovská, I. & Váryová, I. (2017). The Influence of Loans Development on Competitiveness of Agricultural Companies in the Slovak Republic. *Proceedings of the Agrarian Perspectives XXVI: Competitiveness of European Agriculture and Food Sectors*, Prague, Czech Republic, 174-181.

Navarro-Esbrí, J., Diamadopoulos, E. & Ginestar, D. (2002). Time series analysis and forecasting techniques for municipal solid waste management. *Resources, Conservation and Recycling*, 35(3), 201–214. https://doi.org/10.1016/S0921-3449(02)00002-2

Pechrová, M. & Šimpach, O. (2017). Modelling the Development of Eggs Prices in the CR. *Proceedings of Hradec Economic Days, vol 7(1)*, Hradec Králové, Czech Republic, 682–688.

Sabu, K. M. & Kumar M. (2020). Predictive analytics in Agriculture: Forecasting prices of Arecanuts in Kerala, *Procedia Computer Science*, 171, 699–708. https://doi.org/10.1016/j.procs.2020.04.076.

Toth, M., Rabek, T. & Strapekova, Z. (2020). Impact of Integration and Globalization on Business Risk and loans in Slovak Agriculture. *Proceedings of the 19th International Scientific Conference Globalization and its Socio-Economic Consequences 2019 - Sustainability in the Global-Knowledge Economy*, Rajecké Teplice, Slovakia, 05027.

Yuan, M. C. & Yang, J.Z. (2018). Resolving Difficulties of Agriculture-related Small and Micro Enterprises in Applying for Loans. *Proceedings of 4th International Conference on Economics, Management and Humanities Science (ECOMHS 2018)*, South Korea, 40-44.

Contact

Ondřej Šimpach

Prague University of Economics and Business, Faculty of Informatics and Statistics, Department of Statistics and Probability

W. Churchill Sq. 1938/4, 130 67 Prague 3

ondrej.simpach@vse.cz