Central European Review of Economic Issues
EKONOMICKÁ REVUE

Final look at GDP forecasting by Czech institutions

Jiří ŠINDELÁŘ[a*]

[a] *University of Finance and Administration, Prague, Estonská 500, 101 00 Praha 10, Czech Republic*

Abstract

This paper deals with the evaluation of Czech institutions' (the Ministry of Finance and the Czech National Bank) real GDP growth forecasting performance between 1995 and 2015. Contrary to the author's previous papers on this topic, the set-up was altered, in order to assess an 18-month-long annual prediction and set a third estimate as the real-time data input. Using a battery of three error measures (MAE, RMSE, MASE) augmented by the Wilcoxon and Kruskal–Wallis tests, we have found that the MF and the CNB forecasts do not contain a systemic bias. Also, despite some isolated performance deficiencies (i.e. during the recession periods), the accuracy of forecasts prepared by both the MF and the CNB does not differ significantly from the benchmark forecasts of international institutions. Our outcomes hence correspond with the results of previous studies, implying that the changed data set-up does not affect the predictive accuracy of both institutions.

Keywords

accuracy measures; Czech National Bank; dynamic stochastic general equilibrium model; GDP forecasting; Ministry of Finance; subjective adjustments.

JEL Classification: E37, E66, H68, O47

---

*sindelar@mail.vsfs.cz (corresponding author)
The author specifically wishes to thank the Macroeconomic Predictions Unit at the MF, for highly valuable feedback on the original papers and input for further analysis. This article was created with the contribution of institutional support for long-term conceptual development of the research organization University of Finance and Administration.

ISSN 1212-3951 (Print), 1805-9481 (Online)
ER-CEREI, Volume 24: 5–12 (2021).
doi: 10.7327/cerei.2021.03.01

Final look at GDP forecasting by Czech institutions

Jiří ŠINDELÁŘ

1. Introduction

Forecasting of the country's gross domestic product (GDP) development remains a volatile point of concurrent research. Particularly predictions produced by (supra)national bodies, such as finance ministries or central banks, are commonly reviewed both internally (e.g. Keereman, 1999; Daníelsson, 2008) and externally (Öller and Barot, 2000; Allan, 2013). Such attention is understandable, given the importance those predictions represent both in fiscal and monetary policies. As evinced by Frankel's (2011) study, errors in GDP growth forecast significantly influence a country's budgetary results, particularly when overoptimism bias is present, leading to notable deficits. Effects in the business world can be presumed as similarly important (Jaimovich and Rebelo, 2009).

In Central-Eastern Europe (CEE), contrary to the Western situation, comprehensive evaluation of institutional forecasting performance is mostly missing. Specifically in the Czech Republic, most papers focus on both specific settings, such as short horizon (Arnoštová et al., 2011) and a very limited timeline (Antal et al., 2008; Antoničová et al., 2009). Or they employ troublesome methodological apparatus, such as percentage error measures, when the values oscillate around zero (Novotný and Raková, 2011), the Diebold–Mariano test with timelines exhibiting high serial persistency[1] (Vacková, 2014) and using methodology that does not capture variance in forecast error and is sensitive to outliers[2] (MF, 2013). Given the similarity of the forecasting models the Ministry of Finance, Czech Republic (MF) and the Czech National Bank (CNB) use,[3] this creates an important opportunity for detailed analysis. Such analysis should be performed in the context of new machine-learning techniques (Rajkumar, 2017; Richardson et al., 2018), which arguably provide an important potential for forecast accuracy improvement.

Chronologically, this paper expands on the author's preceding papers on GDP growth evaluation (Šindelář, 2017; Šindelář and Budinský, 2016) with altered methodology and a different data set-up. The goal of this paper is to evaluate the accuracy of real GDP growth annual forecasts produced by Czech central institutions (MF, CNB) in the period of 1995–2015. In order to reach this goal, a two-step approach was adopted: (i) first a set of accuracy measures ranging from scale-dependent to scaled errors was calculated for quantitative comparison. In the second step (ii) we used a battery of tests (Kruskal–Wallis test, Wilcoxon signed ranked t.) to determine the most common performance traits, such as systemic bias or mutual differences. Results of both parts are then summarised and discussed, predominantly with the results of both the aforementioned papers of the author.

2. Data

Our database is formed by the total number of 21 annual real GDP growth forecasts produced by the MF and the CNB between 1995 and 2015. We utilise summer predictions produced, mostly published in July of the year preceding the year being forecast, as our forecast value (Ft). Such a setting implies an 18-month (18M) horizon as being evaluated. On the other side, our real value (Yt) is composed of the summer value presented in the OECD Economic outlook in year+2 after the forecast is created (early out-turn).[4] With respect to the previous, we have utilised the CZSO (2018) and the OECD (2018) as our main data sources.

[1] Well documented by Christensen et al. (2007), the Diebold–Mariano test exhibits substantial problems in dealing with finite time-series and serial persistence (rejecting null too often – oversized type I error), making it unsuitable.

[2] As examined by Makridakis and Hibon (1995) and Hyndman and Koehler (2006), used range of error measures does not capture variance in forecast error (average forecasting error, mean average error) and is sensitive to outliers (Theil´s Inequality Coefficient – TIC), particularly because of using RMSE as TIC's relative measure.

[3] Both MF (Alieyev et al., 2014) and CNB (Andrle et al., 2009) utilise the expanded dynamic stochastic general equilibrium model (DSGE) adjusted by expert judgements of forecasting staff.

[4] Published usually in June–July, this means that we compare e.g. July 2010 forecast of 2011 GDP growth with GDP growth data for 2011 published in July 2012.

Both F_t and Y_t data parts represent expansion of the author's previous papers, which utilised different time horizons (3M, 9M, 15M and 21M) and first (Šindelář, 2017) and most recent (Šindelář and Budinský, 2016) out-turn, respectively. This new setting not only sheds light on the most important budgetary forecasting horizon (government budget is drafted using 18M growth horizon), but also brings the analysis in line with benchmark papers like Öller and Barot (2000) or Daníelsson (2008). Because of this, the current paper closes the final evaluation gap and completes the analytical circle.

3. Method

Apart from the data changes, the methodology remains the same as in the original papers. That is, we use a battery of three forecasting errors to evaluate the forecasts (let us denote the forecasting error E_t as the difference $Y_t - F_t$):

- Mean Absolute Error (MAE)

$$MAE = mean(|E_t|).$$

- Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{mean(E_t{}^2)}.$$

- Mean Average Scaled Error (MASE)[5]

$$MASE = mean \frac{E_t}{\frac{1}{n-1}\sum_{i=2}^{n}|Y_i - Y_{i-1}|}.$$

Using these three measures, and particularly MASE instead of TIC error, we follow the guidelines set by prolific papers in the field, such as Hyndman and Koehler (2006) or Armstrong and Collopy (1992). With the three measures above, we are able to cover all the crucial aspects of forecasting performance, such as the magnitude of forecasting error, systemic bias and performance in changes. This compensates for deficiencies present in other Czech studies, outlined in the introduction part.

In the second step, we have undertaken two separate statistical tests to analyse the significance of selected traits:

- Presence of systemic bias – we used the **Wilcoxon test** as our primary method, augmented by the T-test. Application of the Wilcoxon test for such a purpose is common among forecasters (see e.g. Mühleisen et al., 2005; Campbell and Ghyssels, 1995 or Daníelsson, 2008). In comparison with parametric alternatives, this test does not require the errors to be normally or t-distributed. Its two assumptions of symmetry and independence were pretested using the Box–Pierce independence test and the Miao, Gel and Gastwirth symmetry test, providing favourable results (outlined in appendix no. 1).
- Comparison with benchmark forecasts (OECD, European Commission, consensus forecast) – because of previously mentioned limitations of the Diebold–Mariano test, we utilised the less restrictive non-parametric **Kruskal–Wallis test**. As with the Wilcoxon test, this method demands that the errors are independently drawn from a continuous and symmetric population – both assumptions were not rejected in the previous paragraph.

P-values less than 0.05 were considered statistically significant. Analysis was conducted using the R statistical package, version 3.2.3.

Finally, the paper follows on methodology emulating the learning-test sample division. A similar approach was utilised in one of the previous analyses (Šindelář and Budinský, 2016), when conducting first versus most recent out-turn evaluation. Lack of significant differences between both data sets can be interpreted as an acceptable model fit in terms of forecasting methods used by surveyed institutions, similarly to procedures described by Gareth et al. (2013).

4. Results

Table 1 summarises the forecasting errors we have calculated, in comparison with the author's previous paper (Šindelář, 2017) results.

As with the original paper, the error measures indicate three basic findings. Firstly, both institutions clearly struggle with forecasting turning points and discontinuities, as documented by a dramatic increase of error (MAE, RMSE) in the 2008–2010 and 1996–1998 periods. The growth eras, such as 1999–2002 and most recently 2014–2015, exhibit a much better performance. This confirms that turning points remain a crucial forecasting challenge, which can greatly benefit from adding machine-learning techniques to the traditional dynamic stochastic general equilibrium (DSGE) model, adjusted by expert judgement. Neural networks, a premier machine-learning technique, were reported to be the only method able to forecast *surprises* (i.e.

[5] In the original paper, MASE was calculated for individual years and then averaged. This approach heavily penalises less accurate forecasts and represents a deviation from the computation algorithm suggested by Hyndman and Koehler (2006). In this paper, we strictly compute MASE on an interval basis, i.e. as a scaling vector over a timeline of forecasts forming total and subperiods.

growth discontinuities)[6] and offer about one-quarter lower forecasting errors when it comes to traditional alternatives, such as autoregression or general equilibrium.[7] This can result in significant improvement of predictive accuracy and, in our context, even *break* the performance of both MF and CNB versus the naïve benchmark, as reported next.

Secondly, the comparison with the naïve benchmark (MASE) remains troublesome for both institutions, but mainly for the ministry. It was able to surpass the naïve forecast only in three periods, one turning (2008–2010) and the other stable (1999–2002 and 2014–2015). The CNB, on the other hand, failed to do so only in the steep growth period (2003–2007), when it consistently undershot the real value. These partial results aggregate to the bank's better than naïve performance for the overall period (0.85), while the ministry exhibited almost the same accuracy, compared to the naïve benchmark (1.01). Finally, the mutual comparison reveals that the CNB was, on average, able to achieve smaller errors on the new 18M horizon. It needs to be noted, however, that the longer time frame captured by the ministry data (including *surplus* 1996–1998 recession) penalises total MF forecasting performance over the shorter CNB time line in this comparison.

Compared to the original paper, the MAE and RMSE error metrics retained a comparable pattern to the previous forecasting horizons (3M, 9M, 15M and 21M), with the highest values related to the described turning points. As of their amplitude, the 18M forecast represents an almost smooth transition between shorter (15M) and longer (21M) horizons, fulfilling well the general expectation on error horizon proportion. The MASE measure, however, offers a different picture. Switching to strictly interval values, we have found this method of computation to indicate notably smaller error sizes. In this new set-up, we have found that the CNB predictions surpass the naïve benchmark (MASE < 1) and the MF ones are on the verge of doing so, talking about the total period. Much more favourable results were attained in subperiods as well. This sheds a different light on an important part of institutions' forecasting performance, in a positive way that will be discussed later.

5. Statistical tests

As in the original paper, we have used two groups of tests to determine whether systemic bias is present and whether the accuracy of the MF and the CNB forecasts are different from set benchmarks. The outcomes of the first step can be found in Table 2.

As evident from the results, on the selected p = 0.05 level, the systemic bias was overwhelmingly not detected in either the MF or the CNB forecasts (Shapiro–Wilk and Bai–Ng normality tests were utilised to decide which of the two main tests would be used, but nevertheless these provided the same outcome). At this point, therefore, the results are fully compliant with the findings of the previous paper.

Similarly, with the previous table, no differences were found in terms of forecast accuracy between the MF/CNB predictions and the selected benchmarks (consensus, the OECD and the EC forecasts). This implies that none of the surveyed institutions performed better or worse than the rest of the sample in a statistically significant way. Again, this upholds the findings of the original paper in the new, updated set-up.

[6] See Rajkumar (2017) for details.
[7] See Richardson et al. (2018) for details.

Table 1 Error measures – a comparison

Period	Actual paper						Šindelář (2017)					
	MF (18M)			CNB (18M)			MF (18M)[A]			CNB (18M)[A]		
	MAE	RMSE	MASE	MAE	RMSE	MASE	MAE	RMSE	MASE	MAE	RMSE	MASE
1996–2015[B] (total period)	2.4	2.99	1.01	1.9	2.52	0.85	2.350	3.107	3.068	2.200	2.593	3.308
1995–2001	2.6	3.03	1.15	1.9	2.01	1.10	2.350	2.921	1.625	1.850	1.551	1.120
2002–2007	1.9	2.05	1.33	1.7	1.89	1.19	1.800	2.117	2.572	1.850	2.137	2.779
2008–2013	3.0	4.08	0.90	2.4	3.52	0.73	2.700	4.015	5.121	2.650	3.721	5.296
1996–1998 First recession	3.9	4.00	1.32	-	-	-	2.050	2.332	2.756	-	-	-
1999–2002 Recovery	1.1	1.34	0.76	1.4	1.68	0.95	1.250	1.579	0.857	1.700	1.994	1.067
2003–2007 Steep growth	2.2	2.24	1.45	2.0	2.07	1.34	2.150	2.313	3.047	2.050	2.250	3.201
2008–2010 Second recession	3.9	5.27	0.82	3.2	4.54	0.69	3.750	5.271	0.678	3.100	4.643	0.563
2011–2013 Stagnation – third recession	2.1	2.35	1.08	1.6	2.04	0.81	2.000	2.099	9.565	2.400	2.455	10.030
2014–2015 Recovery[A]	1.1	1.42	0.47	1.0	1.12	0.43	-	-	-	-	-	-

[A] The total period in this paper is two years longer than in the original one (which ended in 2013).
[B] Because the original paper evaluated 15M and 21M horizons, we used an approximate 18M result by arithmetically averaging those two.

Table 2 Systemic bias – a comparison

Test	Actual paper				Šindelář (2017)			
	18M Forecast				18M Forecast			
	1995–2015	1995–2001	2002–2007	2008–2015	1995–2015	1995–2001	2002–2007	2008–2013
MF_Wilcoxon test	0.466	0.219	0.313	0.641	0.5885	0.305	0.1875	0.313
MF_T-test	0.294	0.141	0.306	0.414	0.3325	0.286	0.179	0.2575
MF_Shapiro-Wilk Normality t.	0.068	0.605	0.307	0.150	0.1635	0.808	0.782	0.524
MF_Bai-Ng Normality t.	0.292	0.447	0.235	0.017	0.2385			
CNB_Wilcoxon test	0.794	0.625	0.438	0.400	0.727	0.6875	0.1565	0.312
CNB_T-test	0.565	0.717	0.358	0.320	0.743	0.4825	0.203	0.228
CNB_ Shapiro-Wilk Normality t.	0.015	0.314	0.054	0.169	0.076	0.9705	0.329	0.478
CNB_ Bai-Ng Normality t.	0.325	0.062	0.011	0.020	0.308			

Table 3 Differences in accuracy – a comparison

Test	Actual paper				Šindelář (2017)			
	18M Forecast				18M Forecast			
	1995–2013	1995–2001	2002–2007	2008–2013	1995–2013	1995–2001	2002–2007	2008–2013
Kruskal-Wallis test (all together)[8]	0.967	0.915	0.992	0.994	0.910	0.624	0.9795	0.980

6. Discussion and conclusions

The aim of this discussion paper was to amend the evaluation carried out in the previous (Šindelář, 2017) study. Although we used an altered data set-up, by using different real-time data (Y_t) and a forecasting horizon, the results in a strong majority of most cases support the original findings:

- Absolute forecasting errors were found to sharply increase in discontinuity periods connected to macro-economic recessions, as previously observed by Öller and Barot (2000) or Daníelsson (2008).
- Neither the MF nor the CNB forecasts were detected to carry systemic bias of either overforecasting or sandbagging, confirming their internal credibility.
- None of the surveyed institutions produces GDP growth forecasts that are significantly worse (or better) than the other ones on the overall scale, providing external credibility, but raising the question of the institutional value added.

From a factual perspective, special attention should be given to the post-2008 great recession. Contrary to previous crises on a given horizon, this one represented very sharp discontinuity, at least through the optics of surveyed forecasts. Central institutions generally failed to foresee the said discontinuity (Alessi et al., 2014; Christiano et al., 2018) and their Czech counterparts make no exception. Although both of them, the CNB and MF alike, were able to beat the in-sample naïve forecast of the MASE metric, MAE and RMSE errors skyrocketed. Yet, as tested by the author's paper (Šindelář, 2017) in question, performance among other international bodies was comparable (OECD, European Commission). The general (un)predictability of such discontinuities remains a challenge for macro-forecasting. Although some authors speculated about the potential of subjective (expert) methods (Armstrong, 1985), this hypothesis was debunked by our research. All of the surveyed institutions, in fact, use subjective (expert) adjustments as part of their forecasting model. Empirically described overoptimism of economic experts might have contributed to their generally inferior performance (Mathy and Stekler, 2017).

The last point is even more connected to a comparison with the naïve benchmark, evaluated by an adjusted MASE method algorithm. At this point, we have determined that altering the computational method to operate with the scaling (interval) factor instead of averaging separated yearly values has a remarkable impact on measurement results. Error values were strongly reduced and while the MF predictions

[8] Consensus forecast data cover the period of 2000–2015, the OECD data for 1995–2015 and the EC data for 2000–2015.

narrowly remained in the unfavourable zone (providing lower accuracy than naïve in the sample benchmark, while including the additional recession of 1996–1998), the CNB forecasting performance was found to be superior, when speaking of the shorter 1998–2015 period. Methodologically, this development is consistent with Hyndman and Koehler's (2006) recommendations in their baseline paper, which were only partly reflected in the original analysis.

This discussion paper also provides important findings on the side. Altering real-time data (Y_t) from the most recent out-turn to the first out-turn did not have an effect on the study results in terms of the statistical significance of the surveyed traits (systemic bias, benchmark comparison). Such an outcome fully corresponds with observations made in the author's recent work on the topic (Šindelář and Budinský, 2016). Finally, the paper now provides a more comparable basis with regards to already existing dedicated evaluations carried out by the MF (MF, 2013; Vacková, 2014) or the CNB (Arnoštová et al., 2011; Antal et al., 2008; Antoničová et al., 2009; Novotný and Raková, 2011) authors. Still, it paints a more critical picture because of different methods used (MASE, statistical tests), and keeps in place implications on past evaluations' deficiencies (inappropriate use of the MAPE method, problems with the Diebold–Mariano test, among others).

Finally, our results suggest improvement potential connected to machine-learning forecasting. Not only do these new techniques offer potential for further accuracy improvement (Richardson et al., 2018), which – given our results – public institutions struggle to produce over time with traditional methods, but concurrent papers also indicate important value-added when it comes to forecasting surprises and turning points (Rajkumar, 2017). Precisely these discontinuities are the source of the greatest errors with traditional DSGE methods. Our final recommendation, therefore, points to the imminent need for a survey in this promising field and empirical evaluation of said potential. This is the final outcome and also concluding direction for future research.

References

ALLAN, G. (2013). Evaluating the usefulness of forecasts of relative growth, *Strathclyde discussion papers in economics* No. 12/2013.

ALESSI, L., GHYSELS, E., ONORANTE, L., PEACH, R., POTTER, S. (2014). Central bank macroeconomic forecasting during the global financial crisis: the European central bank and Federal Reserve Bank of New York experiences, *Journal of Business & Economic Statistics 32*(4): 483-500. https://doi.org/10.1080/07350015.2014.959124

ALIYEV, I., BOBKOVÁ, B., ŠTORK, Z. (2014). Extended DSGE model of the Czech economy. *Ministry of Finance, Czech Republic Working Paper*, No. 1/2014.

ANDRLE, M., HLÉDIK, T., KAMENÍK, O., VLČEK, J. (2009). Implementing the new structural model of the Czech National Bank. *Monetary Department, Czech National Bank Working Paper*, No. 2/2009.

ANTAL, J., HLAVÁČEK, M., HORVATH, R. (2008). Do Central Bank Forecast Errors Contribute to the Missing of Inflation Targets? The Case of the Czech Republic, *Finance a úvěr: Czech Journal of Economics and Finance* 58(09-10): 434-453.

ANTONIČOVÁ, Z., MUSIL, K., RŮŽIČKA, L., VLČEK, J. (2009). Evaluation of the CNB's Forecasts, *Economic Research Bulletin* 7(1): 8-10.

ARMSTRONG, J. S. (1985). *Long-range forecasting*. New York: Wiley and sons.

ARMSTRONG, J. S., COLLOPY, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons, *International Journal of Forecasting* 8(1): 69-80. https://doi.org/10.1016/0169-2070(92)90008-W

ARNOŠTOVÁ, K., HAVRLANT, D., RŮŽIČKA, L., TÓTH, P. (2011). Short-Term Forecasting of Czech Quarterly GDP Using Monthly Indicators, 6: 566-583.

CAMPBELL, B., GHYSELS, E. (1995). Federal budget projections: A nonparametric assessment of bias and efficiency, *The Review of Economics and Statistics* 77(1): 17-31. HTTPS://DOI.ORG/10.2307/2109989

DANÍELSSON, Á. (2008). Accuracy in forecasting macroeconomic variables in Iceland. *The Central Bank of Iceland Working Paper*, No. 39.

FRANKEL, J. (2011). Over-optimism in forecasts by official budget agencies and its implications, *Oxford Review of Economic Policy*, 27(4): 536-562. https://doi.org/10.1093/oxrep/grr025

GARETH, J., WITTEN, D., HASTIE, T., TIBSHIRANI, R. (2013). *An introduction to statistical learning*. New York, USA: Springer Verlag.

HYNDMAN, R. J., KOEHLER, A. B. (2006). Another look at measures of forecast accuracy, *International Journal of Forecasting*, 22(4): 679-688. https://doi.org/10.1016/j.ijforecast.2006.03.001

CHRISTENSEN, J. H., DIEBOLD, F. X., RUDEBUSCH, G., STRASSER, G. (2007). Multivariate Comparisons of Predictive Accuracy. *University of Pennsylvania, USA Working Paper*, No. 8/2007.

CHRISTIANO, L. J., EICHENBAUM, M. S., TRABANDT, M. (2018). On DSGE models, *Journal of Economic Perspectives*, *32*(3): 113-40. https://doi.org/10.3386/w24811

JAIMOVICH, N., REBELO, S. (2009). Can news about the future drive the business cycle?, *American Economic Review*, 99(4): 1097-1118.

https://doi.org/10.1257/aer.99.4.1097
KEEREMAN, F. (1999). The track record of the Commission forecasts. *Directorate General Economic and Monetary Affairs (DG ECFIN) European Commission Working Paper,* No. 137.
MAKRIDAKIS, S., HIBON, M. (1995). Evaluating Accuracy (or Error) Measures. *INSEAD, France, Working Paper,* No. 95/18/TM.
MATHY, G., STEKLER, H. (2017). Expectations and forecasting during the Great Depression: Real-time evidence from the business press, *Journal of Macroeconomics* 53: 1-15.
https://doi.org/10.1016/j.jmacro.2017.05.006
MÜHLEISEN, M., DANNINGER, S., HAUNER, D., KRAJNYÁK, K., SUTTON, B. (2005). How do Canadian budget forecasts compare with those of other industrial countries? *IMF Working Papers,* 66(5): 1-49.
https://doi.org/10.5089/9781451860856.001
NOVOTNÝ, F., RAKOVÁ, M. (2011). Assessment of Consensus Forecasts Accuracy: The Czech National Bank Perspective, *Finance a Uver: Czech Journal of Economics and Finance* 61(4): 348-366.
ÖLLER, L. E., BAROT, B. (2000). The accuracy of European growth and inflation forecasts, *International Journal of Forecasting* 16(3): 293-315.
https://doi.org/10.1016/S0169-2070(00)00044-3
RAJKUMAR, V. (2017). Predicting surprises to GDP: a comparison of econometric and machine learning techniques. *Doctoral dissertation, Massachusetts Institute of Technology*.
RICHARDSON, A., MULDER, T. (2018). Nowcasting New Zealand GDP using machine learning algorithms. *Reserve bank of New Zealand, NZ. CAMA Working Paper*, No. 47/2018.

Appendix 1 Independence and symmetry test results

Forecast	**Box-Pierce independence test**	**Miao, Gel and Garswith symmetry t.**
MF_18M	0.33	0.468
CNB_18M	0.698	0.328
OECD_18M	0.248	0.51
EC_18M	0.745	0.37
Consensus_18M	0.54	0.402

https://doi.org/10.2139/ssrn.3256578
ŠINDELÁŘ, J. (2017). GDP Forecasting by Czech Institutions: An Empirical Evaluation, *Prague Economic Papers* 2017(2): 155-169.
https://doi.org/10.18267/j.pep.601
ŠINDELÁŘ, J., BUDINSKÝ, P. (2016). Evaluation of gdp growth forecasts: does using different data vintages matter?,*Ekonomický časopis (Journal of Economics),* 9(64): 827-846.
VACKOVÁ, P. (2014). Evaluation of the Ministry of Finance’s Forecast History, *Statistics and Economy Journal* 94(2): 18-35.

Additional sources

MINISTRY OF FINANCE, CZECH REPUBLIC. (2013). *Makroekonomické predikce na MF ČR – pohled do zpětného zrcátka* [online], accessed at 28.5.2018. Available from: <http://www.mfcr.cz/assets/cs/media/Makro-ekonomicka-predikce_2013-Q3_Makroekonomicke-predikce-na-MF-CR-pohled-do-zpetneho-zrcatka-cervenec-2013.pdf>
ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. (2018). *OECD Economic Outlook Archive.* [online], accessed at 28.8.2018. Available from:
<https://stats.oecd.org/index.aspx?queryid=51396>
CZECH STATISTICAL OFFICE. (2016). *HDP, národní účty.* [online], accessed at 28.8.2018. Available from:
<https://www.czso.cz/csu/czso/hdp_narodni_ucty>