Real-time versus revised Czech data: a DSGE analysis

Jan Čapek[1]

Abstract. The paper investigates the effects of using real-time data instead of revised final data within Dynamic Stochastic General Equilibrium (DSGE) framework. Theoretically, using most recent, updated historical data for ex-post analyses for a historical time-sample may be misleading, because revised data were not available at that time.

The contribution uses a small-scale monetary macroeconomic DSGE model to analyze the importance of real-time data compared to most-recent revised data with a focus on the differences in decision-making of the monetary authority. The analysis proceeds from Bayesian estimation of model parameters in a model with real-time data and in a model with the most recent revised data.

Keywords: real-time data, data revision, recursive estimate, DSGE model

JEL classification: C11, C32, C52, C54, E52, F41
AMS classification: 91B51, 91B64, 91B84, 62P20

1 Introduction

The paper investigates the effects of using real-time data instead of revised final data within Dynamic Stochastic General Equilibrium (DSGE) framework. Theoretically, using most recent, updated historical data for ex-post analyses for a historical time-sample may be misleading, because revised data were not available at that time. In case of e.g. monetary macroeconomic models with the usage of revised data, the monetary authority would be expected to make decisions based on then-unknown data. The contribution uses a small-scale monetary macroeconomic DSGE model to analyze the importance of real-time data compared to most-recent revised data with a focus on the differences in decision-making of the monetary authority. The analysis proceeds from Bayesian estimation of model parameters in a model with real-time data and in a model with the most recent revised data. Statistical significance of the differences in the Bayesian estimates is presented.

Following section introduces the reader to real-time data literature. Literature review focuses on papers that are relevant for this paper's aim. Section 3 presents analysis of the data itself with descriptive statistic. Section 4 presents results of recursive estimation on a SOE DSGE model with real-time and revised data. Final section concludes.

The term "real-time data" addresses data that become available right after collection. In accordance with the literature on similar topics, this paper understands "real-time data" as data that are available 3-4 months after the end of a quarter and are typically the first estimates published for that quarter.

A "vintage" is a quarter at which the data becomes available, or, the time of publishing. For example, if Czech statistical office releases during April of 2013 an estimate of GDP growth for last quarter of 2012, it is said that the data for fourth quarter of 2012 are in vintage of April of 2013.

2 Literature review

The issue of data revising and properties of real-time versus revised data has been investigated in the literature for a long time and until now, it is still a topic for scientific discussion.

The orientation of various literature and methodology differs. Some articles focus on the data itself and investigate properties of real-time data in comparison with revised data. There is also a group of articles that focus on the influence of real-time data on model results, e.g. for forecasting or reactions of monetary policy. This contribution follows both of these lines of research.

---

[1]Masaryk University, Faculty of Economics and Administration, Department of Economics, Lipová 41a, 600 02 Brno, capek@econ.muni.cz

Croushore and Stark [2] investigate properties of a real-time dataset and show that data vintages matter in comparison to revised data. Taylor [8] focuses on forecast quality when real-time data are used for its construction. He shows that forecasting models that use timely data have higher forecasting power. Business cycle analysis in real-time data environment is done e.g. by Lee et al. [4], who construct real-time macroeconomic database for Australia. The database exposes the difficulties in drawing inferences and decision-making based on macroeconomic data that is subsequently revised. Molodtsova et al [6] estimate Taylor rule with revised data and real-time data. There is a difference between the two estimates in Germany but not so much in the USA. Orphanides and van Norden [7] examine the reliability of alternative output detrending methods with special attention to the accuracy of real-time estimates of the output gap. They find out that the revision of published data is not the primary source of revisions in measured output gaps; the bulk of the problem is due to the pervasive unreliability of end-of-sample estimates of the trend in output. Boivin [1] searches for the differences in drifting parameters with real-time data. The findings suggest important but gradual changes in the Taylor rule coefficients, not adequately captured by the usual split-sample estimation.

This contribution follows the trends in literature and it presents analysis on the data itself by descriptive statistics in section 3 and it also calculates model results with recursive estimates in section 4.

3   Data analysis

There is a number of sources of uncertainty concerning the data and its model implementation. This contribution addresses three of them: (i) the differences in the data by source, (ii) the magnitude of data revisions and (iii) the differences in detrended data due to data availability at time of detrending.

3.1   Differences in the data by source

Figure 1 depicts the differences between the data from two widely used databases: OECD and Eurostat. The data should be the same with the exception of inflation, which is only available as CPI in OECD real-time database and HICP in Eurostat database. In most cases, the differences are small or none at all, which is the case of both domestic and foreign interest rates with no error and EuroArea growth with only 0.02 percentage points of average absolute error. On the other hand, average absolute difference in data series for Czech inflation differs by a half of a percentage point, which is not negligible. As the errorbars display, there are no systematic patterns in differences between Eurostat and OECD data.
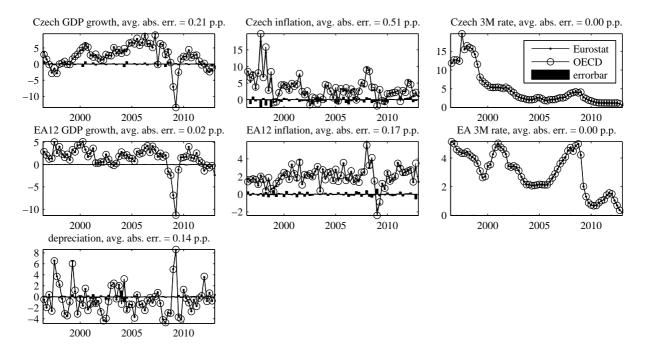

Figure 1 Eurostat and OECD data, *Note*: (error)bar plot around zero = difference between Eurostat and OECD data, "avg. abs. err." = average absolute error in percentage points

3.2 Magnitude of data revisions

Figure 2 displays the series of real-time data and most-recent updated data, all from OECD database. The differences in the series are not insignificant and the errorbars in cases of GDP growths show that the deviations are also non-random. Real-time data tend to underestimate the actual development. For example, in periods of expansion, real-time data underestimate positive growth in the Czech Republic in 2004/2005 and also in EuroArea in 2007/2008. As for recessions, both real-time values for domestic and EuroArea economy underestimate the severity of the crisis of 2009. The average absolute error tops at 1.34 percentage poins (quarterly, per annum) for Czech GDP growth, which means that the average value of the error is approximately 56% of average GDP growth over the sample. *Average error-to-average value* is also very high for EuroArea growth with 39%.
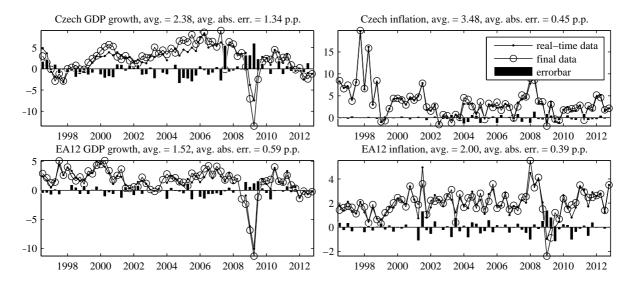

Figure 2 Real-time and final (revised) data, *Note*: (error)bar plot around zero = difference between final and revised data, "avg." = simple average of final data over the sample, "avg. abs. err." = average absolute error in percentage points

3.3 Differences in detrended data due to data availability at time of detrending

This section offers a depiction of a source of uncertainty which is not in the data itself but stems from the need to detrend the time series prior to modeling. In the model used, EuroArea interest rate is detrended with a linear trend and since Czech interest rates display even more curvature, Hodrick Prescott trend is used.

This type of analysis uses only revised data series. At first, the trend is computed on time series truncated to 30 observations. Then, the truncation moves one-by-one to more observations and the trend is recomputed each time. This procedure results in a number of time series of different length, each one detrended on its sample. Such results depict changes in the time series stemming only from the recomputation of the trend.

Figure 3 displays all vintages of detrended data with a solid line and final data with a circled line for comparison. The last values of each vintage form a real-time data in the second row panels. Time series formed this way will be addressed "detrend real-time" data in the remainder of the paper. Note that the difference between "detrend real-time" data and final revised detrended data stems only from the fact that detrending is calculated repeatedly for each length of the series. The third row in Figure 3 displays errorbars that are expectedly systematic.

4 Estimation

4.1 Model and Identification

This paper uses a New Keynesian (NK) Dynamic Stochastic General Equilibrium (DSGE) model. The model is derived from microeconomic behavior of particular economic agents. These include domestic and foreign households, domestic and foreign producers, domestic importers and domestic and foreign monetary authority. The
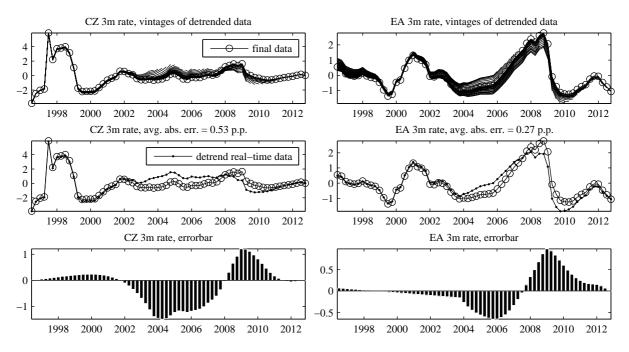
Figure 3 Differences in detrended data due to data availability

model is in small open economy (SOE) setting, so that it presumes two countries - small open economy that is influenced by a big closed economy. The small open economy is the home Czech economy, the big large economy is the foreign EuroArea economy. Most of the model assumptions are adopted from Lubik and Schorfheide [5].

The model is estimated by Bayesian methods in Dynare package[1] on a dataset with 7 variables. All the estimates share the same prior setting. Convergence of each estimate is checked by Brooks and Gelman convergence diagnostics.

4.2 Recursive estimates

Recursive estimates are constructed so that as a new data point is added to the existing set, the model is re-estimated. New values of parameter estimates are then reported. One can argue that if parameters estimates change, it reflects the information in the newly-added data.

Recursive analysis is regularly done by policy institutions if newly-data becomes available to actualize the model for a new round of predictions or forecasting. Recursive analysis is also done on historical sample to analyze the evolution of policy. The former inevitably uses real-time data, the latter mostly uses updated final data, because the estimation is easier to implement and older vintages of real-time data may not be available. This sections shows two examples where the choice of real-time versus final data does matter.

Figures in this section present four estimates to compare. Estimates based on Eurostat and OECD data are standard recursive estimates – only the time series of most updated final data is used, detrended on the whole sample and the detrended series is then truncated to mimic shorter sample. Note that the detrended series do not change at all in this type of estimation, there is just one new data point added in each longer time-frame.

Another approach is a genuine real-time estimate. This sequence of estimates use only data available at the vintage in question. Each time series may undergo revisions and therefore may be different. Also, detrending is made on these potentially different series with different number of observations. Although this approach mimics the historical possibilities of analysis best, it has a drawback that the influence of the data revisions and the influence of potentially changing trends cannot be distinguished.

To address this problem, another recursive estimate is introduced. The original data series for estimation are revised final series, but the truncation of the series is done prior to detrending. This estimation therefore does not reflect any issues of data revisions but it captures the influence of re-detrending in each quarter. It is then possible

---

[1]`www.dynare.org`

to infer the relevance of data revisions by comparing the genuine real-time estimate and the estimate that covers the problems arising from consecutive detrending. Note that this type of results was earlier addressed to as based on "detrend real-time" data. In order to easily see the significance of the results, a 90 % probability band of the "detrend real-time" estimate is drawn.

The estimates start in second quarter of 1996 and have at least 30 observations. The last observation is third or fourth quarter of 2012 which makes 66 or 67 observations in the most-recent estimates.

Estimates based on Eurostat and OECD data usually behave similarly, although there may be some vertical difference between the series of the estimates. However, on the two presented cases, the "detrend real-time" estimate that covers only the problem of consecutive detrending is different than the former two. This means that knowledge of the trend is significant and not knowing future trend leads to different estimates. Also, the real-time estimates differs from "detrend real-time" estimates, which again means that also the nature of the changes in the data itself due to revisions are significant.
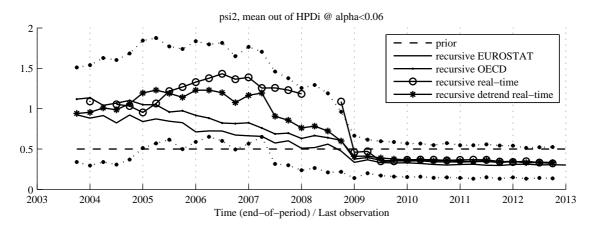

Figure 4 Recursive estimates of the weight on output growth in a domestic Talyor rule

Figure 4 displays the evolution of recursive estimates of a parameter that is the weight on output growth in a domestic Talyor rule. It is therefore a preference parameter of domestic monetary authority. The higher it is, the more important is output growth for the central bank when it set its policy interest rates.

All of the estimates display a tendency of a decline. The tendency is gradual in standard recursive estimates. On the other hand, the real-time estimate shows no gradual tendency but there is a great drop in the estimate in fourth quarter of 2008 (denoted as 2009 in the figure).[2] Since the evolution of the real-time estimate is outside of 90 % probability bands of "detrend real-time" estimate, the difference is statistically significant. Such result may be interpreted intuitivelly – from the point of view of real-time data, the oncoming crisis was a shock that is captured by a major change in the estimate in the quarter in question. Other estimates display gradual decline, because the information may have been included in revised data and the future trend was known.

Figure 5 displays estimation results for the persistence of domestic interest rate and it displays rather different development. Both standard recursive estimates calculated on final revised data tend upwards from 0.6 to final 0.7. However, both remaining estimates also show a decline in years 2006 and 2007 before converging upwards. This results may mean that 2006/2007 period is consistent with lower interest rate smoothing that would seem from the estimates on revised data. Since the evolution of real-time and "detrend real-time" data is similar, the differences are probably mainly due to re-estimation of the trend rather than in data revision itself. Note also that the results are not significant at conventional levels so that we can only discuss tendencies.

5   Conclusion

Data from OECD and Eurostat are very similar and the use in a DSGE model does not make much difference. Comparing revised and real-time data shows that the differences are not insignificant and also with systematic patterns. Also, if the preparation of the data for a model use means detrending (other than demeaning), this makes rise to a problem of trend recomputation, which can also be significant.

---

[2]Note that missing values of the estimates are due to missing real-time data at respective quarters.
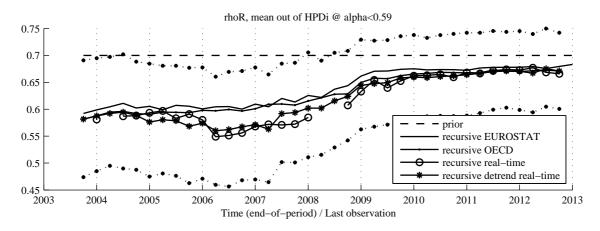
rhoR, mean out of HPDi @ alpha<0.59

Figure 5 Recursive estimates of the persistence of domestic interest rate

Analysis on a DSGE model shows results where the re-detrending is significant and also where the nature of real-time data in comparison to revised data is significant.

Paper's results are in line with existing literature that finds the use of real-time data significant in various estimation or forecasting exercises. This contribution confirms the results on Czech data.

Acknowledgements

References

[1] Boivin, J.: Has U.S. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data. *Journal of Money, Credit and Banking* 38, 5 (August 2006): 1149-1174.

[2] Croushore, D. and Stark, T.: A Real-Time Data Set for Macroeconomists: Does the Data Vintage Matter? *Review of Economics and Statistics* 2003 85:3, 605-617, ISSN 0034-6535

[3] Lee, K., Olekalns, N., Shields, K., and Wang, Z.: Australian Real-Time Database: An Overview and an Illustration of its Use in Business Cycle Analysis, *Economic Record*, Volume 88, Issue 283, pages 495516, December 2012

[4] Lubik, T., and Schorfheide, F.: A Bayesian Look at the New Open Economy Macroeconomics. In *NBER Macroeconomics Annual 2005, Volume 20*, NBER Chapters. National Bureau of Economic Research, Inc, May 2006, pp. 313–382.

[5] Molodtsova, T., Nikolsko-Rzhevskyy, A., and Papell, D. H.: Taylor rules with real-time data: A tale of two countries and one exchange rate, *Journal of Monetary Economics*, Volume 55, Supplement, October 2008, Pages S63-S79, ISSN 0304-3932.

[6] Orphanides, A., and van Norden, S.: The Unreliability of Output-Gap Estimates in Real Time, *Review of Economics and Statistics* 2002 84:4, 569-583, ISSN 0034-6535

[7] Taylor, N.: Economic forecast quality: information timeliness and data vintage effects, *Empirical Economics*, 2013, pp. 1-30, issn 0377-7332.