Department of Applied Econometrics Working Papers

Working Paper No. 2-13

End-of-sample vs. real time data: perspectives for analysis of expectations

Emilia Tomczyk
Warsaw School of Economics

End-of-sample vs. real time data:
perspectives for analysis of expectations

Emilia Tomczyk

Abstract

Data revision is defined as an adjustment published after the initial announcement had been made; it may reflect rectification of errors, availability of new information, etc. When economists use a database, they may not even be aware that some of the values have been revised, perhaps repeatedly, and corrected numbers may significantly differ from original ones. I propose to test whether including information on data revisions helps to model properties of expectations, improve quantification procedures, or adjust tests of rationality to data vintage.

This paper presents review of literature and databases available for the purposes of real time analysis, and offers an introduction to empirical analysis of influence of data vintage on tests of expectations.

1.   Introduction

Several formal methods of analyzing expectations are available, among them descriptive statistics, ex post forecast errors, tests based on frequency tables, entropy measures, and tests of properties of expectations time series. The latter typically consist of testing properties associated with Muth rationality: unbiasedness and orthogonality with respect to available information. To supplement and improve testing procedures, I propose to address the issue of data revisions, that is, corrections published after the data has initially became available. Revised values may significantly differ from original values, and there may be several steps between initial and final – that is, unchanged in any later publications – data.

The question neglected in Polish research, and only intermittently addressed in world literature, is how vintage of data influences results of expectations testing. Should initial or final data be used to compare expectations with their realizations observed at a later date? To quantify survey expectations data? To evaluate unbiasedness and orthogonality of expectations time series?

This paper presents review of literature and databases available for the purposes of real time analysis, and offers an introduction to empirical analysis of influence of data vintage on tests of expectations.

2.   Definitions

End-of-sample (EOS) data is usually defined, following Koenig *et al*. (2003), as data from the latest available announcement – that is, from researcher's point of view, final data. It is important to keep in mind that by the time researcher uses a database, some or even most of the variables had been revised many times, reflecting corrections of errors, newly available information, etc. On the other hand, real time values (RTV) are initial values, made available by statistical agency (or other publisher) directly after their collection. To help describe what data is actually available to economic agents when their make their decisions (or express expectations, or declare observed changes), the term "vintage" is

used; Kozicki (2004) defines vintage of a data series as "the date when that version of the data became available" (p. 8).

To complicate matters, the term "final data" is far from clear. Stark and Croushore (2002) propose the following interpretations of "current values": the latest available; or the last before a benchmark revision; or one year after the observation date. Generally, the researcher can never be sure that the variables he or she is using will never again be revised.

Generally speaking, revisions introduced by statistical agencies are meant to reflect the arrival of new information which became available only after the initial announcement has been made, or correct errors that had been made earlier. Specifically, McKenzie (2006, p. 7) lists the following reasons for revisions of official statistics:

1. Incorporation of data with more complete or otherwise better reporting (e.g. including late respondents) in subsequent estimates.
2. Correction of errors in source data and computations.
3. Replacement of first estimates derived from incomplete samples when more accurate data become available.
4. Incorporation of source data that more closely match the concepts and/or benchmarking.
5. Incorporation of updated seasonal factors.
6. Updating of the base period of constant price estimates.
7. Changes in statistical methodology, concepts, definitions, and classifications.

Jacobs and van Norden (2010, p. 1) add one more reason to the above list: revisions to national accounts statistics based on analysis of supply and demand (input-output) tables.

Points 5 to 7 on McKenzie's list draw attention to the fact that revisions can also result from reasons external to economic content of variables, such as change of the definition of a variable, or weighting procedures used in its construction, or similar issues. Revisions of this type, pertaining to definitions or statistical procedures and not the economic data themselves, are sometimes called spurious revisions. Generally, efforts are made by statistical agencies to remove effects of spurious revisions from the published data, and make the newly published data set compatible with the previous ones. Both extent and efficiency of these efforts are, however, difficult to determine.

The question of EOS vs. RTV data can be placed in the wider context of news-or-noise framework first introduced by Mankiw *et al*. (1984) and Mankiw, Shapiro (1986). They describe "noise" as opposed to "news" and define the latter as efficient messages, incorporating all available and relevant information. For example, Mankiw *et al*. (1984) test whether preliminary announcements of money stock are better described by news hypothesis (that is, initial data are rational forecasts of the final values, and revisions are caused only by new data becoming available) or noise hypothesis (that is, initial announcements are observations of final, or revised, values but are measured with error). This strand of research is continued in recent literature (see for example Borağan Arouba, 2008, and McKenzie *et al*., 2008).

Jacobs and van Norden (2010) classify data revisions into an even wider field of measurement errors. In addition to "noise" component (which they define in terms of uncorrelated measurement errors from different data vintages) and "news" component (defined as characterized by measurement errors equal to rational forecast errors), they propose to consider a "spillover" component in which measurement errors within a given data vintage are serially correlated. They also point out that data revisions may be of seasonal nature, corresponding to data frequency or reflecting publication horizon (for example, once a year).

Typically researchers make every effort to include final (revised) values – that is, EOS data – in their analyses of economic variables. Only recently the consequences and benefits of employing RTV data has been addressed and analyzed, the obvious advantages being that real time economic data enables researchers to reproduce each others' research, even in case of projects based on old and since updated data, and to evaluate policy decisions using the data available at the time.

Koenig *et al*. (2003) emphasize that recent data are usually the least reliable, and before RTV data becomes EOS data, it may undergo several consecutive changes. Hence the need for analysis of their respective merits in analyses of expectations.

3.   Review of literature

Data revisions have been analyzed extensively, if only recently. Even though Jacobs and van Norden (2010) look as far back as 1919, to W. M. Persons' paper published in "The Review of Economics and Statistics", they see the real beginning of real time data analysis in comparative analyses of GDP dating back to the 1950-ties and 1960-ties. The publication usually credited with launching this branch of economic research is the 1955 article by A. J. Gartaganis and A. S. Goldberger published in "Econometrica" (see Croushore, 2011). Since then, the field has expanded and produced numerous research papers on comparisons of forecasts built on the basis of real-time (initial) data and the latest available (final) data, macroeconomic research (including fiscal and monetary policy), and current analysis of business and financial conditions. Various authors have reviewed this literature from different points of view (see Croushore 2006, 2011; Borağan Arouba, 2008; Clements and Galvão, 2010; Cimadomo, 2011) and offered a variety of classifications of the major themes within this field. In this section, I make an attempt to summarize these major topics.

3.1   Data revisions in macroeconomic forecasting

This strand of literature focuses on effects of data vintage on specification of econometric models and evaluation of forecast errors. Researchers attempt to asses whether use of EOS data causes overestimation of predictive value of explanatory variables, or influences forecasts quality, as compared with use of RTV values (keeping in mind that only RTV data were available when forecasts were initially formed). While a lot of papers addressing these issues were published (see Diebold and Rudebusch, 1991; Orphanides, 2001; Croushore and Stark, 2001, 2003; Faust *et al*., 2003; Orphanides and van Norden, 2002; Clements and Galvão, 2010), evidence remains mixed. For example, Diebold and Rudebush find more predictive power in revised than in real time data, but Clements and Galvão find advantages of real time data as compared to end-of-sample values in terms of minimizing the expected squared forecast error. Koenig *et al*. (2003) discuss efficiency of strategies of including different vintages of data; they conclude that use of real time data can overestimate the forecasting power of a model relative to alternative models.

Croushore (2006) presents a review of literature, and summarizes results of his own research, on data revisions and optimal forecasts. He offers the following general suggestions:

- forecasts based on EOS and RTV data differ, and predictive content of variables may change as the result of data revisions,
- forecasts of level variables are revised more often than forecasts of growth rates,
- model choice is influenced by data revisions,
- number of lags in ARIMA-class models is influenced by choice of data vintage.

Influence of data vintage on macroeconomic forecasting seems therefore to be substantial.

A related issue concerns proper response to data revisions when forecasting in real time. It is unclear if using multiple data vintages improves accuracy of forecasts; Croushore (2006) does not find an improvement in including (as opposed to ignoring) data revisions; Clements and Galvão (2010), using VAR methodology, do find advantages of RTV over EOS data, and in a later publication (2011) show that historical (RTV) data supplements EOS values for the purpose of real-time policy analysis. The authors see their results as the support for employment of multiple-vintage models.

Structural approach has also been used for the purpose of multi-vintage analysis of forecasting models. For example, Vázquez *et al*. (2012) propose an extended version of the basic New Keynesian model which includes revision processes of output and inflation data in order to assess the importance of data revisions on monetary policy and on transmission of policy shocks. They find that even though the initial announcements of output and inflation are not rational forecasts of revised data, "ignoring the presence of non well-behaved revision processes may not be a serious drawback in the analysis of monetary policy in this framework" (p. 29).

3.2   Predictability and significance of data revisions

This line of literature addresses the following question: how large and how systematic are data revisions in key macroeconomic variables? Beginning from the 1980-ties, publications focus on variables such as money stock and GDP (see Mankiw *et al*., 1984; Mankiw and Shapiro, 1986; Mork, 1987). The majority of authors find revisions both large and predictable but there is some disagreement on whether they have a significant effect on

estimated monetary policy rules. Croushore (2006, 2011) presents evidence that some data revisions, even after discounting seasonal adjustments, are systematic and predictable and may be a factor in expectations analysis. These findings suggest that institutions which publish and revise economic data can introduce adjustments in the original (preliminary) values in expectations of future revisions, in order to minimize the extent of subsequent updates. However, while revisions in some variables are systematic and predictable, they are unpredictable or very small in others; generally, Croushore shows that revisions do influence quality of forecasts. Borağan Aruoba (2008) confirms these results; he finds initial announcements made by statistical agencies biased, and revisions – large compared to initial data. He concludes that revisions are predictable with information available at the time of the initial announcement. Cimadomo (2011) shows that fiscal data revisions are large, biased and predictable, and that different fiscal policies are suggested by use of real time data as compared to use of end-of-sample data. A summary of previous research on this subject is presented in Arnold (2012); on the basis of earlier publications and her own research, she concludes that whether announcements are revised systematically, and whether economic agents aim at forecasting final or initial values, influences evaluations of forecast accuracy, and finds revisions considerable and at least partly predictable.

Empirical analysis has been centered on, but not limited to, United States. Faust *et al*. (2005) confirm predictability of revisions in GDP growth rates in G-7 countries; in several of them they find updates large and predictable. Golinelli and Parigi (2007) evaluate forecasting performance of preliminary releases of GDP growth for various vintages of US and Italian data. To summarize, data revisions are generally found to be large and, at least to some extent, predictable – although whether they are significant for policy making purposes remains to be seen.

3.3   Frequency and regularity of data revisions

This strand of literature addresses the crucial question of data revisions: when a value can be considered final? In general, revisions may be irregular (for example, benchmark revisions) or regular (for example, with frequency compatible with publication cycle or seasonality patterns). Jacobs and van Norden (2010) propose a framework to distinguish regular revisions from "surprise" ones, and first revisions from subsequent updates. They conclude that measurement errors introduced by data revisions are characterized by

complex dynamics which were not included in previous models, and propose methods to differentiate between "first" and "later" revisions. As far as benchmark revisions are concerned, Kozicki (2004) finds them important for monetary policy purposes, and Phillips and Nordlund (2012) show that there is a cyclical and seasonal bias in the annual benchmark revisions in employment data. In general, both regular and irregular revisions seem to influence results of econometric modeling, and including revisions in a research procedure is far from straightforward.

3.4   Policy making

There are numerous papers on influence of data revisions on monetary policy; for example, Orphanides (2001) shows that policy measures based on initial data – that is, values available at the time policy decisions were made – would be different if undertaken on the basis of revised data. Use of real-time data in analyses of fiscal policy has also been extensively studied (see Forni and Momigliano, 2005; Golinelli and Momigliano, 2006; Bernoth *et al*., 2008; Cimadomo, 2008; and Beetsma *et al*., 2009, published under the enticing title "Planning to cheat: EU fiscal policy in real time"). The vast literature has been reviewed in Cimadomo (2011), and its major findings were summarized as follows: revisions in fiscal data tend to be large and systematic; strong fiscal rules promote accurate reporting and smaller revisions; and "(…) the ex-ante reaction of fiscal policies to the economic cycle is estimated to be more 'counter-cyclical' when real-time data are used instead of ex-post data" (p. 30). To summarize, monetary and fiscal policy decisions are sensitive to vintage of data used for their evaluation.

3.5   Modeling data revisions with state-space models

This strand of literature groups technical papers on forecasting revised data in real time using the Kalman filter (see Conrad and Corrado, 1979) and, later, more advanced methods, among them non-linear and non-Gaussian filters (see Mariano and Tanizaki, 1995). Related papers address the topic of data vintage in presence of measurement errors (see Harrison *et al*., 2005; Jacobs and van Norden, 2010) and modeling multivariate data revisions in systems of variables with linear state space models (see Patterson, 2003; Croushore, 2006; Jacobs *et al*., 2010). State-space models are often used for the purpose of vintage data analysis; when revision process in presented in state-space form, standard

filtering techniques can be used for estimation, inference, forecasting, and imputation of missing data. Perhaps the most comprehensive publication is the one by Jacobs and van Norden (2010). They place updated series within the measurement error models, and propose a space-state framework designed to model a wide set of measurement errors, including data revisions.

4. Sources of data on revisions

Until recently, analysis of data revisions required painstaking work on matching values published in different paper publications in many points of time. In the past two decades, databases containing both real-time data and subsequently revised time series became available. Accessibility of these data sets makes it possible to assess the extent of data revisions.

The most comprehensive real time database is the OECD data set which in addition to all OECD countries also includes the Euro zone countries, China, India, Brazil, South Africa, and the Russian Federation. The Web interface[1] allows access to data for 21 economic variables as originally published in each monthly edition of the Main Economic Indicators database from February 1999 as well as the revisions made to initially published data. They key economic variables include Gross Domestic Product and its expenditure components, industrial production, production in construction, balance of payments, composite leading indicators, consumer prices, retail trade, unemployment rates, civilian employment, hourly earnings, monetary aggregates and international trade values. Time series dating back to the 1960-ties are provided for some variables.

The declared purpose of the database is to provide originally published data for researchers interested in testing performance (for example, forecasting performance) of econometric models in simulated real-time, and to provide data for studies of influence of data revisions, for example analyses of magnitude and direction of revisions to published statistics.

---

[1] Database available at http://stats.oecd.org/mei/default.asp?rev=1

Another database, widely used in applied research, is the Federal Reserve Bank of Philadelphia Real-Time Data Set (RTDS).[2] Introduced in Croushore and Stark (2001), and credited with generating current interest in the analysis of influence of data vintage in economics, it is extensively employed for macroeconomic analysis of effects of data revisions and for analysis based on real-time data (see Croushore and Stark, 2003; Croushore, 2006). It provides the information set that would be available to a forecaster on a 15th day of the middle month in every quarter, starting in 1965 and covering quarterly data on, among other variables, real and nominal output, consumption, investment, and price and employment series. Its limitation is that it covers US data only.

Another branch of the Fed system, Federal Reserve Bank of St. Louis, also publishes vintage data. Its ALFRED database[3] (ArchivaL Federal Reserve Economic Data, also known as Economic Data Time Travel) provides vintage versions of economic data that were available on specific dates in history. The database currently covers 65,037 series in 9 categories, with the earliest vintage for Industrial Production Index being 1927. This data set is also limited to US series.

Vintage data for European economies is published by The Euro Area Business Cycle Network (EABCN). This organization seeks to provide an interface for policy makers from Central Banks and other central institutions, including the European Central Bank, and academic researchers. Its Real Time Database[4] (RTDB) covers data for the euro area and other European countries, and includes over 200 macroeconomic time series of different vintages, acquired from ECB's Monthly Bulletin reports.

Less comprehensive sources are sometimes used, for example vintage data on labor productivity published by the Bureau of Labor Statistics (see Borağan Arouba, 2008). Generally, there seem to be abundant sources of economic real time data – however, the most complete of them are limited to United States series.

---

[2] Database available at http://www.phil.frb.org/econ/forecast/reaindex.html.
[3] Database available at http://alfred.stlouisfed.org/
[4] Database available at http://www.eabcn.org/; detailed description published in Giannone *et al*. (2010).

5. Data revisions in Poland

To the best of my knowledge, there is no Polish database dedicated to collecting real time economic data. Poland is included, however, in the OECD data set described in the previous section. Searching this database proves that some of the data remain virtually unchanged in subsequent publications. For example, from September 2010 till May 2011, data on civilian employment in Poland for the period of third quarter of 2002 – fourth quarter of 2010 have not changed.[5] Similarly, in the same period data on Consumer Price Index (measured as the average changes in the prices of consumer goods and services purchased by Polish households with 2005 level = 100) values for January 2008 – February 2011, published from September 2010 to May 2011, have not changed. On the other hand, data on business conditions (namely, composite leading indicator) exhibit changes across dataset. From July 2011 till April 2012, values for January 2008 – December 2011 have changed significantly in many cases.

For the variables not included in the OECD data set, it is necessary to collect historical data from individual paper or electronic publications. Publications of the Central Statistical Office (CSO) only report the most recent version and if any revisions of earlier data are introduced, the researchers have to identify them themselves. Polish Central Statistical Office (CSO) occasionally publishes notes on revisions. They generally result from conforming with ESA 1995 (The European System of Accounts, most recently updated in 1995), recommendations of Directorate-General of the European Commission (Eurostat), continuing efforts to improve quality of statistics, and new legal documents of UE; see CSO 2007). These revisions include:

- revision of national accounts for the years 1995 – 2004 (in 2005),
- revision of national accounts; regional accounts were re-calculated on the basis of revised data (in 2007),
- revision of public deficit and central and local government debt for the years 2005 – 2008; they were caused by change of methods of valuating income taxes, and including transactions relating to public-private partnerships (in 2009).

National Bank of Poland also publishes occasional announcements on data revisions. For example, on June 29th, 2011, National Bank of Poland declared continuing negative

---

[5] Database accessed in August 2012.

balance of errors and omissions and, in cooperation with the IMF, revised its balance of payments for years 2004 – 2010. Possible revisions in Polish GDP and other macroeconomic series are discussed in financial press and internet forums. For example, on March 22nd, 2013, The Wall Street Journal has quoted the Ministry of Finance economist, Ludwik Kotecki, as saying that "economic expansion in Poland is likely to slow to 1.5%-2% this year, putting budget revenue under pressure and potentially forcing the government to revise its budget deficit goals".[6] Nevertheless, announcements of National Bank of Poland and Ministry of Finance do not provide amount of data necessary for systematic analysis of data revisions and their influence on the Polish economy.

To summarize, access to initial or preliminary (and subsequently revised) Polish economic data is difficult, and constitutes a challenge for empirical analysis of effects of data revisions. Any attempt to analyze Polish real time data, and to evaluate influence of data vintage on behavior of economic time series, either must be based on data available in the OECD database, or will require building a specialized data set on the basis of numerous Central Statistical Office publications. The latter option, though time-consuming, appears promising; for example, Polish general business conditions index undergoes significant revisions. Forecasts of general business conditions in November 2007 differ by 4 percentage points between March 2008 and November 2008 (see Tomczyk, 2011, p. 161). Such noticeable revisions offer an opportunity to test whether they influence other economic variables, and whether data vintage has a significant impact on the results of empirical analysis of Polish time series.

6. Data revisions and testing of expectations

So far, neither extent of data revisions nor their influence on quantification procedures or evaluating predictive properties of expectations were analyzed in Poland, and are only rarely addressed in world literature. Arnold (2012), on the basis of extrapolative and adaptive models of expectations formation build for individual expert (professional forecasters) data, finds that there are no significant differences in expectations formation

---

[6] "Polish Budget Stressed as Growth to Slow", by P. Wasilewski
(http://blogs.wsj.com/emergingeurope/tag/ludwik-kotecki/)

processes for the latest revision and for the initial values. As far as I am aware, aggregated expectations time series were not analyzed from the data vintage point of view.

I believe that three main issues arise when data is revised in context of analysis of expectations.

First, in regard to process of expectations formation, which data is used by the economic agents to formulate expectations: end-of-sample (that is, final) or data available in real time, or perhaps some intermediate values published between the first and final announcements? Do economic agents expect data revisions when forming their expectations? Do they consider real time data to be reliable, and do EOS and RTV series differ with respect to reliability?

Second, when evaluating quality of expectations and their accuracy with respect to observed values, should RTV or EOS data be used? Which of these types of data should be employed to assess expectations (forecast) errors?

Third, should RTV or EOS data be used for the purpose of quantification of survey data on expectations? Quantification methods (both probabilistic and regression-based) require that survey data on values observed by respondents be compared with "official" quantitative data series – but should they be end-of-sample or real time data?

Two general suggestions were offered in my earlier work (see Tomczyk, 2011):
- When designing quantification models, survey data should be compared with final (EOS) data. Respondents are probably aiming to describe their real observations and not initial guesses subject to adjustments.
- When evaluating accuracy of expectations of economic agents, particularly whether all information has been employed (orthogonality), RTV data (that is, available at the moment of expectations formation) should be used, even if they were later corrected. Expectations should not be evaluated by comparison with final values because economic agents, at the time of their formation, did not have access to revised data, or to information that would enable them to assess the extent of revisions (unless they are predictable which has not been proved for Polish data).

As the review of literature presented in Section 3 shows, there is no shortage of literature on effects of data vintage on modeling and forecasting economic phenomena. However, the majority of papers is based on highly aggregated macroeconomic variables like GDP and money stock, and analyze data revision issues in the framework of fiscal or monetary policy. One of the reasons for such a focus may be importance of fiscal and monetary policy design and evaluation for practical purposes. Another reason, however, may be that key macroeconomic data is subject to major and systematic data revisions due to the high level of aggregation and numerous difficulties in reporting. Those factors cause initial data to be frequently revised, and open a window of opportunity for researchers interested in data revisions. Less aggregated variables – for example, monthly or quarterly series on industrial production, prices, or employment – are very rarely, if ever, analyzed for issues associated with data vintage. I would like to propose to extend analysis of data revisions to less aggregated data, keeping in mind that specialized datasets will have to be constructed for this purpose.

As the second sub-field of data revisions analysis I would like to suggest the following topic: whether – and if yes, how strongly – results of quantification procedures and rationality tests are influenced by data vintage. Tests of rationality of expectations in Poland have failed to provide conclusive results, and neglecting data vintage issues may be one of the reasons for lack of unambiguous conclusions.

To summarize, I would like to propose that effects of data vintage be analyzed from the point of view of properties of expectations series as explanatory variables in econometric models, quantification procedures employed for survey data, and rationality of expectations. Instead of focusing on highly aggregated macroeconomic variables, less aggregated monthly data should be used for this purpose. There are several reasons for this choice of dataset, among them longer time series, avoiding the problem of different sampling frequency for various variables, and possibility of comparative analysis, since previous research on properties of expectations in Poland was based, in significant part, on monthly data.

There is one additional advantage of employing monthly – and, consequently, relatively long – time series. Availability of a large sample enables the researcher to test properties of expectations in subsamples, and draw conclusions about, for example, sensitivity of

expectations formation processes or quantification procedures to external shocks or phases of economic cycles. In the data revision framework, it is possible that behavior of variables in later (that is, based on more current data) subsamples differs from early subsamples because of difficulties associated with introducing adjustments in quality of goods and services (so-called hedonic adjustment). Landefeld and Grimm (2000) show that close to 18% of GDP in the United States is deflated by hedonic measures, that is, quality-adjusted prices are used. Since hedonic adjustments to already published data can be classified as data revisions, another field of research opens for economists interested in effects of data vintage on modeling economic processes.

7. Final remarks

The purpose of this paper was twofold: to present a review of literature and databases available for the purposes of real time analysis, and to propose a framework for future research of effects of data vintage on properties of expectations. While empirical analysis of influence of data revisions on expectations time series is currently under way, there are three related projects which are also worth attention:

- analysis of predictability of data revisions in Poland,
- determining a number of periods after which there is no further (non-spurious) revisions to data series; this number may differ with respect to type and frequency of data,
- developing real time datasets with Polish macroeconomic data for the purposes of further analysis of data vintage.

I believe that analysis of data revisions may help describe properties of expectations and improve quantification procedures and rationality tests, and therefore enhance our understanding of behavior of economic processes.

8. References

Arnold E. A. (2012) *The role of revisions and uncertainty in professional forecasts*, paper presented at the National Bank of Poland Workshop "Are We Really Forward-Looking? Measuring and Testing Expectations – Central Bank Perspective", Warszawa

Beetsma R., Giuliodori M., Wierts P. (2009) *Planning to cheat: EU fiscal policy in real time*, Economic Policy 24:753-804

Bernoth K., Hughes Hallett A., Lewis J. (2008) *Did Fiscal Policy Makers Know What They Were Doing? Reassessing Fiscal Policy with Real Time Data*, The Centre for Economic Policy Research Discussion Paper No 6758

Borağan Aruoba S. (2008) *Data revisions are not well-behaved*, Journal of Money, Credit and Banking 40:319-340

Cimadomo J. (2008) *Fiscal Policy in Real Time*, ECB Working Paper Series No 919

Cimadomo J. (2011) *Real-time data and fiscal policy analysis: a survey of the literature*, ECB Working Paper Series No 1408

Clements M. P., Galvão A. B. (2010) *Real-time forecasting of inflation and output growth in the presence of data revisions*, presented at the Federal Reserve Bank of Philadelphia Conference, 18-19.10.2010, Philadelphia

Clements M. P., Galvão A. B. (2011) *Improving Real-time Estimates of Output Gaps and Inflation Trends with Multiple-vintage Models*, Department of Economic Working Paper No. 678, Queen Mary University of London

Conrad W., Corrado C. (1979) *Application of the Kalman filter to revisions in monthly retail sales estimates*, Journal of Economic Dynamics and Control 1:177-198

Croushore D. (2006) *Forecasting with real-time macroeconomic data*, w: G. Elliott, C. W. J. Granger, A. Timmermann (red.) *Handbook of Economic Forecasting. Volume I*, Elsevier B.V., Amsterdam, s. 961-982

Croushore D. (2011). *Frontiers of real-time data analysis*, Journal of Economic Literature 49:72-100

Croushore D., Stark T. (2001) *A real-time data set for macroeconomists*, Journal of Econometrics 105:111-130

Croushore D., Stark T. (2003) *A real-time data set for macroeconomists: Does the data vintage matter?*, The Review of Economics and Statistics 85:605-617

CSO (2007) *Gross Domestic Product. Regional Accounts in 2005*, Central Statistical Office and Statistical Office in Katowice, Katowice

Diebold F. X., Rudebusch G. D. (1991) *Forecasting output with the composite leading index: A real-time analysis*, Journal of the American Statistical Association 86:603-610

Faust J., Rogers J. H., Wright J. H. (2003) *Exchange rate forecasting: The errors we've really made*, Journal of International Economic Review 60:35-39

Faust J., Rogers J. H., Wright J. H. (2005) *News and noise in G-7 GDP announcements*, Journal of Money, Credit and Banking 37:403-419

Forni L., Momigliano S. (2005) *Cyclical sensitivity of fiscal policies based on real-time data*, Applied Economics Quarterly 50:299-326

Gartaganis A. J., Goldberger A. S. (1955) *A note on the statistical discrepancy in the national accounts*, Econometrica 23:166-173

Giannone D., Henry J., Lalik M., Modugno M. (2010) *An Area Wide Real Time Database for the Euro Area*, ECB Working Paper No 1145

Golinelli R., Momigliano S. (2006) *Real-time determinants of fiscal policies in the Euro area*, Journal of Policy Modeling 28:943-964

Gollinelli R., Parigi G. (2007) *GDP forecasting with real-time data*, presentation at the 27th Annual International Symposium on Forecasting, New York

Harrison R., Kapetanios G., Yates T. (2005) *Forecasting with measurement errors in dynamic models*, International Journal of Forecasting 21:595-607

Jacobs J. P. A. M., Sturm J.-E., van Norden S. (2010) *Modeling multivariate data revisions*, presentation at the 30th CIRET Conference in New York

Jacobs J. P. A. M., van Norden S. (2010) *Modeling data revisions: Measurement error and the dynamics of „true" values*, „Journal of Econometrics" (doi:10.1026/j.jeconom. 2010.04.010)

Koenig E. F., Dolmas S., Piger J. (2003) *The use and abuse of real-time data in economic forecasting*, The Review of Economic and Statistics 85:618-628

Kozicki S. (2004) *How do data revisions affect the evalutaion and conduct of monetary policy?*, Federal Reserve Bank of Kansas City Economic Review, First Quarter 2004

Landefeld J. S., Grimm B. T. (2000) *A note on the impact of hedonics and computers on real GDP*, Survey of Current Business, December 2000, p. 17–22

Mankiw N. G., Runkle D. E., Shapiro M. D. (1984) *Are preliminary announcements of the money stock rational forecasts?*, Journal of Monetary Economics 14:15-27

Mankiw N. G., Shapiro M. D. (1986) *News or noise: An analysis of GNP revisions*, Survey of Current Business 66:20-25

Mariano R. S., Tanizaki H. (1995) *Prediction of final data with use of preliminary and/or revised data*, Journal of Forecasting 14:351-380

McCracken M. W., Clark T. E. (2007) *Tests of equal predictive ability with real-time data*, presentation at the 27th Annual International Symposium on Forecasting, New York

McKenzie R. (2006) *Undertaking revisions and real-time data analysis using the OECD main economic indicators original release data and revisions*, OECD Statistics Working Paper STD/DOC(2006)2; http://search.oecd.org/ officialdocuments/displaydocumentpdf/ ?doclanguage= en&cote=std/doc(2006)2

McKenzie R., Tosetto E., Fixler D. (2008) *Assessing the efficiency of early release estimates of economic statistics*, Working Paper, Organisation for Economic Co-operation and Development, Paris

Mork K. A. (1987) *Ain't behavin': Forecast errors and measurement errors in early GNP estimates*, Journal of Business and Economic Statistics 5:165-175

Orphanides A. (2001) *Monetary policy rules based on real-time data*, American Economic Review 91:964-985

Orphanides A., van Norden S. (2002) *The unreliability of output gap estimates in real time*, Review of Economic and Statistics 84:569-583

Patterson K. D. (2003) *Exploiting information in vintages of time-series data*, International Journal of Forecasting 19:177-197

Persons W. M. (1919) *Indices of business conditions*, Review of Economic Statistics 1:5-107

Phillips K. R., Nordlund J. (2012) *The efficiency of benchmark revisions to the current employment statistics (CES) data*, Economic Letters 115:431-434

Stark T., Croushore D. (2002) *Forecasting with a real-time data set for macroeconomists*, Journal of Macroeconomics 24:507-531

Tomczyk E. (2011) *Oczekiwania w ekonomii. Idea, pomiar, analiza* [*Expectations in Economic. Definitions, Measurement, Analysis*], Oficyna Wydawnicza SGH, Warszawa

Vázquez J., María-Dolores M., Londoño J. M. (2012) *The effect of data revisions on the basic New Keynesian model*, International Review of Economics and Finance, doi: 10.1016/j.iref.2012.03.005