DALLASFED

Economic Letter

Real-Time Historical Dataset Enhances Accuracy of Economic Analyses

by Adriana Z. Fernandez, Evan F. Koenig and Alex Nikolsko-Rzhevskyy

> *A growing body of empirical macroeconomic literature suggests that analyses using real-time data often yield substantially different—and more accurate—conclusions.*

Revised U.S. gross domestic product (GDP) growth numbers released in summer 2011 revealed that the national economy was in worse shape two years after the recession ended than earlier data had suggested and that the downturn itself had been deeper than previously estimated. Revisions such as these from government agencies are commonly issued to account for errors, data updates and measurement changes. Such adjustments involve important economic variables and affect not only the latest available statistic, but also the historical properties of an entire data series. That means revisions can be far-reaching, affecting structural model results, forecasts and monetary policy.

When data are subject to change, *real-time data*—the information available to researchers and policymakers at the time they conducted their analyses—rather than the most up-to-date figures are necessary to appropriately assess a particular economic model or forecast or understand a given monetary policy action. A growing body of empirical macroeconomic literature suggests that analyses using real-time data often yield substantially different—and more accurate—conclusions than those relying on the final revisions.[1]
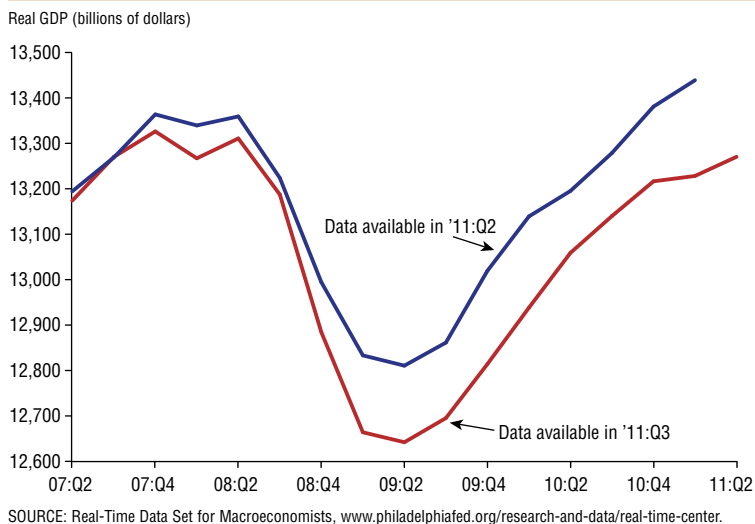
However, such research remains limited, largely reflecting the difficulties of compiling real-time data and the technical complexity of using "vintages," or snapshots of data at points in time.

Economists Dean Croushore and Tom Stark published their large Real-Time Data Set for Macroeconomists (RTDSM) roughly a decade ago, with snapshots of the U.S. economy starting in 1965. Their work established the importance of real-time data and became the U.S. dataset for forecasters and others engaged in research affected by data revisions. Yet very little work has been done to collect and analyze such figures for economies outside the U.S.—even as globalization has made real-time international data increasingly relevant. The Original Release Data and Revisions Database (ORDRD) from the Organization for Economic Cooperation and Development (OECD) is the most comprehensive and well-maintained international real-time database. Updated monthly, it provides vintages of monthly and quarterly data for member countries beginning in January 1999.[2] Its drawback: It covers only the past decade, even though the OECD's recorded figures go back to the organization's inception in 1961.

Seeing the value in extending the dataset, Federal Reserve Bank of Dallas

Chart 1
U.S. Real GDP Undergoes Major Revision in Summer 2011

Real GDP (billions of dollars)

SOURCE: Real-Time Data Set for Macroeconomists, www.philadelphiafed.org/research-and-data/real-time-center.

researchers took on the task of compiling a comprehensive quarterly dataset of 13 variables for each of the 26 OECD countries for which sufficient data were available, drawing from hard copies of historical documents and the OECD Main Economic Indicators from 1962 through 1998. The result of that work, the Real-Time Historical Dataset for the OECD (RTHD-OECD), is a complementary real-time dataset that can be easily merged into the existing OECD real-time dataset.[3] A current version of RTHD-OECD can be downloaded at www.dallasfed.org/institute/oecd/index.cfm. (While our dataset went through many checks to ensure quality, minor errors may still exist. Therefore, a preliminary release has been made available for comment at www.rthd-oecd.org.)

Working with 'Vintages'

The downward revision to real GDP in the U.S. in summer 2011 revealed that the pace of inflation-adjusted economic growth had substantially decreased during the second quarter. Revisions to prior quarters showed real output in the 2007–09 recession fell much more than initially estimated (*Chart 1*). The chart demonstrates how the picture can change when new vintages of data are used.

If we move a few vintages back, the U.S. real GDP real-time data series would resemble the matrix in Table 1, where

each successive column characterizes the snapshot of quarterly data containing the information available at that vintage date.[4] Third quarter 2011 revisions reveal a slower economy not only in second quarter 2011 but dating back to the beginning of the recession. Traditional revised-data research would use only the last column of data—or the latest available information—ignoring previous revisions.

An important aspect of real-time research is analysis of revisions. Generally, when governments make efficient use of all available information, revisions add "news" and are not predictable between vintages. When they don't, revisions merely reduce "noise" and are inefficient, and later values may be predicted.[5]

The RTHD-OECD presents the opportunity to assess efficiency in a longer-vintage span than previous real-time data allowed. Categorizing revisions by the lag length (in quarters) with which they are released, we look for recognizable patterns that would provide evidence against efficiency—specifically whether revisions are significantly positive or negative over the whole historical period.[6]

At a minimum, efficiency requires that revisions to a series should be zero on average. Accordingly, we check whether revisions differ significantly from zero. The results show that in 16 of 26 countries, revisions seem to be predictable (i.e., inef-

ficient) for at least one of the four variables considered (*Table 2*).[7]

The positive readings in virtually all significant revisions suggest that statistical agencies may have a tendency to underestimate inflation and growth in real GNP/GDP, the price level, industrial production and money supply in their earlier estimates. Looking at the overall dimension of corrections, the absolute value of mean revision analysis (*Chart 2*) suggests that in international research, corrections are simply too large to be ignored—as traditional revised-data research does.

Four Important Applications

The RTHD-OECD may be used in many areas of international macroeconomic research in which data revisions matter. Four important applications illustrate the potential of the dataset—which, when merged with ORDRD, provides coverage from first quarter 1962 to second quarter 2010:[8]

1. Testing some of the most frequently used output-gap estimation techniques. The output gap, a key statistic in many important macroeconomic models, shows the difference between an economy's potential and current output.[9] By assessing output-gap estimation methods, researchers and policymakers can identify those generating the most accurate signals.

2. Assessing the predictive ability of the output gap. The output gap is often used as an indicator of future inflation, based on the empirically observed relationship between the two variables (initially identified by economist A.W. Phillips and known as the "Phillips curve").[10] Using the combined dataset, we find that the additional predictive power of the output gap is minimal or nonexistent in real time.

3. More accurately gauging the effect of inflation when interpreting data revisions. In most theoretical models, inflation is thought to have a significant, but temporary, impact on the economy. With the dataset, we find that by making accounting more difficult, higher inflation amplifies the causes and extent of data revisions, which could increase the likelihood of policy mistakes.[11]

4. Spotting vulnerabilities in nominal exchange-rate forecasting models. Most of these models are developed and tested

Table 1	How U.S. GDP Data Change During Successive Revisions

Vintages (billions of dollars)

Period	09:Q2	09:Q3	09:Q4	10:Q1	10:Q2	10:Q3	10:Q4	11:Q1	11:Q2	11:Q3
2007:Q3	11,625.70	13,321.10	13,321.10	13,321.10	13,321.10	13,268.50	13,268.50	13,268.50	13,268.50	13,269.80
2007:Q4	11,620.70	13,391.20	13,391.20	13,391.20	13,391.20	13,363.50	13,363.50	13,363.50	13,363.50	13,326.00
2008:Q1	11,646.00	13,366.90	13,366.90	13,366.90	13,366.90	13,339.20	13,339.20	13,339.20	13,339.20	13,266.80
2008:Q2	11,727.40	13,415.30	13,415.30	13,415.30	13,415.30	13,359.00	13,359.00	13,359.00	13,359.00	13,310.50
2008:Q3	11,712.40	13,324.60	13,324.60	13,324.60	13,324.60	13,223.50	13,223.50	13,223.50	13,223.50	13,186.90
2008:Q4	11,522.10	13,141.90	13,141.90	13,141.90	13,141.90	12,993.70	12,993.70	12,993.70	12,993.70	12,883.50
2009:Q1	11,340.90	12,925.40	12,925.40	12,925.40	12,925.40	12,832.60	12,832.60	12,832.60	12,832.60	12,663.20
2009:Q2	n.a.	12,892.40	12,901.50	12,901.50	12,901.50	12,810.00	12,810.00	12,810.00	12,810.00	12,641.30
2009:Q3	n.a.	n.a.	13,014.00	12,973.00	12,973.00	12,860.80	12,860.80	12,860.80	12,860.80	12,694.50
2009:Q4	n.a.	n.a.	n.a.	13,155.00	13,149.50	13,019.00	13,019.00	13,019.00	13,019.00	12,813.50
2010:Q1	n.a.	n.a.	n.a.	n.a.	13,254.70	13,138.80	13,138.80	13,138.80	13,138.80	12,937.70
2010:Q2	n.a.	n.a.	n.a.	n.a.	n.a.	13,216.50	13,194.90	13,194.90	13,194.90	13,058.50
2010:Q3	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	13,260.70	13,278.50	13,278.50	13,139.60
2010:Q4	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	13,382.60	13,380.70	13,216.10
2011:Q1	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	13,438.80	13,227.90
2011:Q2	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	13,270.10

SOURCE: Real-Time Data Set for Macroeconomists at www.philadelphiafed.org/research-and-data/real-time-center.

Table 2	Significant Average Revisions Denoting 'Inefficient' Revisions

	Real GNP/GDP Release lag				Price level Release lag				Industrial production Release lag				Money supply Release lag			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
U.S.					0.12**				0.28**		−0.19*					
Australia		0.47**		−0.30**			0.28*	−0.29**		−0.85*				−0.27*		
Austria									−0.34**							
Canada		0.28*						0.22*								
Denmark								−0.41*							0.33*	
Finland											−1.17*					
Germany										0.30**	−0.31*			0.19*		
Greece										0.86*					−0.16*	
Italy						0.50*								0.24*		
Japan			0.23**						0.32*							
Mexico			−1.40*	2.21**								1.62**	5.63**			
Netherlands			0.42**													
Norway								0.52*								
Portugal														−0.48*		
Switzerland					1.18**										0.29**	
Turkey					2.86**					1.97**	2.23**					

NOTES: All variables are expressed in terms of annualized quarter-over-quarter growth rates. Significance at 5 and 10 percent is denoted with * and **, respectively. Real gross national product (GNP) is used when real gross domestic product (GDP) is unavailable.

SOURCE: Authors' calculations.

using revised data.[12] Our real-time analysis adds evidence suggesting that revised data analysis may result in misleading conclusions. A particularly interesting case is the British pound, whose exchange rate is predictable at the short horizon with revised data but not with real-time data.

Importance of Real-Time Data

A growing body of empirical macroeconomic literature supports the importance of real-time data analysis. Making use of more-efficient real-time information, researchers can more easily separate news from noise and more accurately detect patterns in data. Given increased globalization and the advantages of using real-time data, international researchers may want to rethink the practice of using revised data because of the potential for misleading conclusions. The RTHD-OECD can serve as a standard for forecasters and others engaged in international research who confront data revisions.

Fernandez is an economist in the Houston Branch and Koenig is a vice president and policy advisor in the Dallas office of the Federal Reserve Bank of Dallas. Nikolsko-Rzhevskyy is an assistant professor of economics at the University of Memphis.

Notes

[1] See "A Real-Time Data Set for Macroeconomists," by Dean Croushore and Tom Stark, *Journal of Econometrics*, vol. 105, no. 1, 2001, pp. 111–30; "Monetary Policy Rules Based on Real-Time Data," by Athanasios Orphanides, *American Economic Review*, vol. 91, no. 4, 2001, pp. 964–85; "Is the Markup a Useful Real-Time Predictor of Inflation?" by Evan Koenig, *Economics Letters*, vol. 80, no. 2, 2003, pp. 261–67; and "Taylor Rules and Real-Time Data: A Tale of Two Countries and One Exchange Rate," by Tanya Molodtsova, Alex Nikolsko-Rzhevskyy and David H. Papell, *Journal of Monetary Economics*, vol. 55, Supplement, 2008, pp. S63–79.
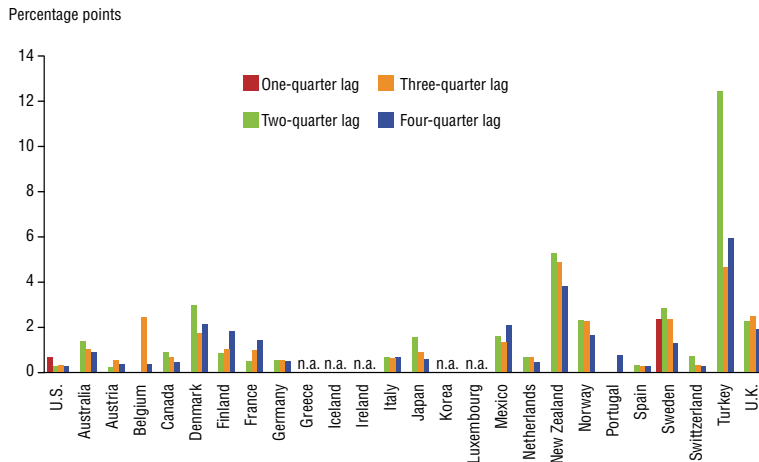
[2] The official OECD Original Release Data and Revisions Database is publicly available at http://stats.oecd.org/mei.

[3] The OECD General Statistics Bulletin was replaced after 1964 by the OECD Main Economic Indicators. See

Chart 2 | Data Revisions Matter Globally

Absolute Mean Revisions for Real GNP/GDP Too Large to Ignore

Percentage points

Legend: One-quarter lag, Two-quarter lag, Three-quarter lag, Four-quarter lag

Countries: U.S., Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, U.K.

Absolute Mean Revisions for Industrial Production Appear Substantial

Percentage points

Legend: One-quarter lag, Two-quarter lag, Three-quarter lag, Four-quarter lag
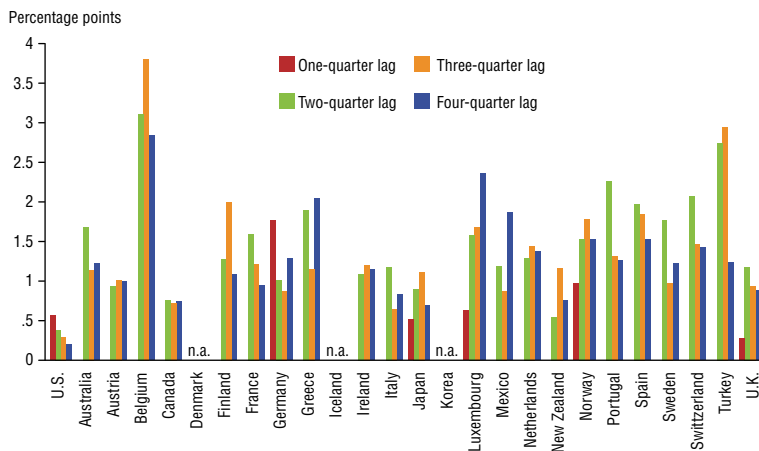
NOTES: Revisions are defined as the difference in the value of a given variable in subsequent vintages. "Absolute mean revision" refers to the total magnitude of changes; that is, adding changes whether they're positive or negative. All variables are expressed in terms of annualized quarter-over-quarter growth rates. Real gross national product (GNP) is used when real gross domestic product (GDP) is unavailable.

SOURCE: Authors' calculations.

www.rthd-oecd.org for the list of countries and variables used in the RTHD-OECD dataset.

[4] Note that real GDP data are released with one quarter lag, which implies that the last data point contained in each vintage corresponds to the previous quarter's real output.

[5] Revisions are not predictable when *all* the available information is incorporated because, in that case, the error term is independent and identically distributed. See "Risk and Return: Consumption Beta Versus Market Beta," by N. Gregory Mankiw and Matthew D. Shapiro, *The Review of Economics and Statistics*, vol. 68, no. 3, 1986, pp. 452–59.

[6] In our dataset, we found revisions released with one through six lags. For this analysis, we present only the results for one- through four-quarter-release lags.

[7] For details on estimation and data employed in subsequent tables and charts, refer to "A Real-Time Historical Database for the OECD," by Adriana Z. Fernandez, Evan F. Koenig and Alex Nikolsko-Rzhevskyy, Federal Reserve Bank of Dallas Globalization and Monetary Policy Institute, Working Paper no. 96, December 2011.

[8] For the empirical applications, we include only the G-7 economies. For details on estimation and data, see note 7.

[9] Potential output refers to the output an economy would produce if all its resources were fully employed.

[10] For the analysis, we follow "The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time," by Athanasios Orphanides and Simon van Norden, *Journal of Money, Credit and Banking*, vol. 37, no. 3, 2005, pp. 583–601.

[11] The case is particularly clear in real GNP/GDP, where we obtained results showing that a 10 percent increase in a country's inflation rate in a given year would increase growth-rate revisions by a full 1 percent over the following year.

[12] See "The Out-of-Sample Failure of Empirical Exchange Rate Models: Sampling Error or Misspecification?" by Richard Meese and Kenneth Rogoff in *Exchange Rates and International Macroeconomics*, Jacob A. Frenkel, ed., Chicago: University of Chicago Press, 2003, pp. 67–112; "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability," by Nelson C. Mark, *American Economic Review*, vol. 85, no. 1, 1995, pp. 201–18; and "Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals," by Tanya Molodtsova and David H. Papell, *Journal of International Economics*, vol. 77, no. 2, 2009, pp. 167–80.