Institute for International Economic Policy Working Paper Series
Elliott School of International Affairs
The George Washington University

Differences in Early GDP Component Estimates Between Recession and Expansion

IIEP-WP-2011-05

Tara M. Sinclair and H.O. Stekler
Department of Economics
George Washington University
Washington DC 20052

February 16, 2011

Institute for International Economic Policy
1957 E St. NW, Suite 502
Voice: (202) 994-5320
Fax: (202) 994-5477
Email: **iiep@gwu.edu**
Web: **www.gwu.edu/~iiep**

Differences in Early GDP Component Estimates Between Recession and Expansions[1]

Tara M. Sinclair[*] and H.O. Stekler[**]
Department of Economics
George Washington University
Washington DC 20052

JEL Codes: C82, E32, C53

Keywords: Flash Estimates, Data Revisions, GDP Components, Statistical Tests, Business Cycles

Abstract

In this paper we examine the quality of the initial estimates of the components of both real and nominal U.S. GDP. We introduce a number of new statistics for measuring the magnitude of changes in the components from the initial estimates available one month after the end of the quarter to the estimates available 3 months after the end of the quarter. We further specifically investigate the potential role of changes in the state of the economy for these changes. Our analysis shows that the early data generally reflected the composition of the changes in GDP that was observed in the later data. Thus, under most circumstances, an analyst could use the early data to obtain a realistic picture of what had happened in the economy in the previous quarter. However, the differences in the composition of the vectors of the two vintages were larger during recessions than in expansions. Unfortunately, it is in those periods when accurate information is most vital for forecasting.

*Sinclair: tsinc@gwu.edu,**Stekler:hstekler@gwu.edu

[1] We would like to thank Deanna Jensen for her outstanding research assistance. We benefited from the comments made on an earlier draft by Daniel Bachman, Ulrich Fritsche, Tom Mayer, Blake Saville, and Xuguang Sheng. We thank Laurent Ferrara for a very helpful discussion of our paper and participants in the 7th Workshop of the International Institute of Forecasters in Verbier, Switzerland. We are grateful to the Institute for International Economic Policy that generously supported this research. We alone are responsible for any remaining errors.

Differences in Early GDP Component Estimates Between Recession and Expansions

Tara M. Sinclair and H.O. Stekler

Over the years, there has been considerable interest in both the financial and academic communities about the nature and extent of the revisions of the GDP data. The financial interest is observed in the way that announcements of data revisions are reported and dissected by the media. Usually the focus has been on the headline numbers, rather than on an extensive analysis of the underlying components of GDP.

The previous academic studies have examined a wide variety of issues: Are the revisions of the US data so substantial that they prevent analysts from correctly interpreting either the state of the economy or the changes that have occurred? (Zellner, 1958; Morgenstern, 1963; Stekler, 1967). Are the early numbers optimal forecasts of the *final* data or are they figures which represent measurement errors, i.e. do the revisions represent news or noise? (Mankiw et al., 1984; Mankiw and Shapiro, 1986; Mork, 1987; de Leeuw, 1990; Neftci and Theodossiou, 1991; Faust et al., 2005; Aruoba, 2008).

There have also been studies which examine the effect that data problems might pose for nowcasting or forecasting.[2] In making predictions about the behavior of the economy, forecasters and policymakers need to know the state of the economy in recent past quarters as reflected in the early or flash estimates. This involves a tradeoff between accuracy and timeliness. How accurate are the early GDP data released 15-30 days after the end of the quarter relative to the revised figures released 30 or 60 days later? Are the early data accurate enough to provide correct information about the state of the economy, especially prior to and during

[2] For issues involved in nowcasting, see Stark and Croushore, 2002; Croushore, 2006, 2009, 2010. There have also been analyses of the real time conduct of monetary and fiscal policies. See Croushore (2009) for studies that have examined this issue; Groen et al. (2009) analyzed the Bank of England's real time forecasts.

recessions? (McNees, 1986; Zarnowitz 1982; Joutz and Stekler, 1998; Dynan and Elmendorf, 2001, Swanson and van Dijk, 2006).

In terms of our current knowledge about the relationship between revisions and real time analysis, the evidence is that the revisions are large and systematic. For example, the growth rate estimate for 1977.1 varied between 4.9% and 9.6% depending on the vintage of the data. (Croushore, 2006). More recently, the early estimates for 2008.4 were between -3.8% and -6.2%. The mean absolute revision in the growth rate of GDP has been around 1%.[3] Dynan and Elmendorf (2001) concluded that "…provisional estimates do not fully capture accelerations and decelerations, suggesting some tendency to miss economic turning points." Similarly, Joutz and Stekler (1998) concluded that while the early data were useful to forecasters, there were three turning point errors in the early data. All of these errors occurred during recessions; two of these happened at the end of the cycles.

These studies have focused on the headline GDP numbers. They did not investigate how the revisions affected the *components* of a particular vintage of GDP data.[4] The main topic of this paper is to determine whether data revisions affect the estimates of the components of GDP substantially from one vintage of data to another. We introduce a number of new statistics for measuring the magnitude of these changes. We further specifically investigate whether the state of the economy can affect these changes.

[3] The revisions are substantial even in evaluating five year average growth rates. (Croushore, 2009).

[4] The US Bureau of Economic Analysis (BEA) that publishes these data has, however, examined the extent of the data revisions in the components that aggregate to GDP (Young, 1987; de Leeuw, 1990; Young, 1993; Grimm and Parker, 1998; Fixler and Grimm, 2002, 2005, 2008). These analyses primarily focused on the differences between the early estimates and the numbers that were available at the time the research was done. These studies usually did not discuss the extent of the revisions between the data that were released approximately 30, 60, and 90 days after the end of the quarter to which they refer. The paper by Fixler and Grimm (2008) is an exception but it only analyses the mean and mean absolute revisions of current dollar GDP. There is no discussion of the revisions to the real variables.

The next section presents the methodology for measuring the bias and the statistics that are used to measure the changes in estimates of the components in the various vintages of data. This is followed by a description of the data and the results. The implications of our results comprise the concluding section.

I. Methodology

As mentioned above, most of the focus in past analyses has been on headline GDP estimates rather than on the estimates of the underlying components of GDP. It matters whether, for example, a change in our estimate of GDP occurred as a result of a change in the estimate of final demand or in inventory accumulation. It may also matter whether a change in the estimate occurred as a result of a change in the estimate of consumption or investment expenditures. We currently do not know whether there is a significant change in the estimates of any of the underlying components of GDP when the data are revised. If GDP revisions are biased, we may learn something about the cause of that bias based on our analysis of the underlying components. Even if GDP estimates are unbiased, there may be interesting offsetting biases in the underlying components. While we first investigate whether the earliest estimates are unbiased estimates of the data that are available two months later, we will specifically examine and focus on the changes in the estimates of the underlying components. We will then determine whether the changes that occur near or during a recession are significantly different from those that occur in expansionary periods. We examine this issue by dividing our sample into expansionary and recessionary quarters.

In terms of the timeliness-accuracy tradeoff, we postulate that the earliest data are valuable if they are unbiased estimates of the later data, if there are no significant compositional *changes between the two sets of estimates and if there are no significant differences between the results for periods of expansion and recession*.

A. Bias

We first investigate whether the earliest nominal and real GDP estimates are unbiased estimates of the final numbers available two months later. We next undertake the same analysis for each of the major components of both nominal and real GDP. Customarily, the basic procedure for testing for bias has been to use the Mincer-Zarnowitz (1969) regression.

$$y_{3t} = \beta_0 + \beta_1 y_{1t} + e_t, \tag{1}$$

where y_{3t} and y_{1t} are the U.S. Bureau of Economic Analysis (BEA)'s estimates for time t available 3 months and 1 month after the quarter, respectively. For a test of informational efficiency, the null hypothesis is: $\beta_0 = 0$ *and* $\beta_1 = 1$. A rejection of this hypothesis indicates that the initial estimates are biased and/or inefficient. The Wald test and the F distribution are used to test this null.

An alternative procedure for testing for bias has been to use equation (2) as suggested by Holden and Peel (1990):

$$y_{3t} - y_{1t} = \beta_0 + e_t. \tag{2}$$

In the Holden and Peel test, the slope is imposed to be one and the test examines whether or not the data revision has a zero mean. Thus the bias test in this case is a simple test of statistical significance for the constant in equation (2).

Recent research has shown that forecasts sometimes contain systematic errors (Joutz and Stekler, 2000, Hanson and Whitehorn, 2006). Forecasters overestimated the rate of growth during slowdowns and recessions and underestimated it during recoveries and booms. Similarly, inflation was under-predicted when it was rising and over-predicted when it was declining. In some cases, these systematic errors, associated with the stages of the business cycle, may offset each other yielding regression estimates that do not reject the null of bias when in fact there are errors that are associated with the state of the economy. In order to determine whether the GDP estimates similarly failed to incorporate information about the state of the economy, we modified (1) as in Sinclair et al. (2010). The modified Mincer-Zarnowitz regression (3) now becomes

$$y_{3t} = \beta_0 + \beta_1 y_{1t} + \beta_2 D_t + e_t, \tag{3}$$

where D_t is a dummy that reflects the state of the economy. It takes on the value 1 if during one month of a particular quarter the economy was in a recession. Otherwise, the value of the dummy is zero. For this calculation, the data for the quarter before the peak and the quarter after the trough were included with the numbers for the quarters that constituted the recession as defined by the National Bureau of Economic Research (NBER).[5] The justification for this procedure comes from Young (1987, p. 29) who considered those to be the most critical quarters from the BEA's perspective. The joint null hypothesis now is: $\beta_0 = 0$, $\beta_1 = 1$, *and* $\beta_2 = 0$. If any of the coefficients associated with the dummies are non-zero, the dummies contain information that can explain the initial estimation errors. If this were to occur, it would indicate that the BEA had not included the information about the state of the economy in the initial estimates. A similar modification can be applied to the Holden and Peel test by adding the same dummy to regression (2).

These tests for the existence of bias were applied to the data for each GDP component separately. It is also possible to determine whether the vector of these GDP components was biased. For this test we estimated a system of equations representing the Mincer-Zarnowitz

[5] The NBER dates are available here: http://www.nber.org/cycles/cyclesmain.html.

regressions for the 10 components of GDP using a seemingly unrelated regression (SUR) model. We created four systems: the Mincer-Zarnowitz regressions for the nominal data, the Mincer-Zarnowitz regressions for the real data, the modified Mincer-Zarnowitz regressions for the nominal data, and the modified Mincer-Zarnowitz regressions for the real data. The SUR model estimates the coefficients allowing for heteroskedasticity and contemporaneous correlation in the errors across equations. For each system we are able to perform a Wald test that $\beta_0 = 0$ and $\beta_1 = 1$ (and $\beta_2 = 0$ for the modified Mincer-Zarnowitz regressions) for all 10 equations simultaneously.

B. Compositional Changes: Difference of Two Vectors

When the BEA releases its National Income and Product Account (NIPA) estimates, it does not just provide the headline numbers, the growth rates of real and nominal GDP. It also releases estimates of each of the major components of GDP. These numbers, which show the growth rates of the components, can be viewed as a vector comprising a particular vintage of data relating to that particular quarter. When all the data for that quarter are subsequently revised, the components of that vintage of data comprise a different vector. Thus, if we are concerned with how well the estimates reflect the actual changes that have occurred in the economy, we must compare the difference in the vectors of the different vintages.[6]

[6] Under some circumstances, the weighted sum of the changes of these elements will add up to the total change in GDP. Under these conditions and if in addition all of the changes are strictly positive, there are methods for comparing the composition of two vintages of data referring to the same time period. Theil (1966) developed an information inaccuracy measure that compared forecasts with outcomes. In our context, this approach would measure the value of the original data given the information in the revised data. It is called the information inaccuracy of the earliest data. The more value that is associated with the newer data, the less valuable were the earlier statistics. Patterson and Heravi (1991) used this method to measure the value of the revisions to the components of UK GDP, but their analysis was in the levels (not changes) of these components. They include the net change in inventories as part of the investment component and also make an adjustment for imports. Patterson and Heravi (2004) applied cointegration tests to the various vintages of the US data as an alternative measure of accuracy for the early data. Öller and Teterukovsky (2007) use a different information concept to analyze the quality of some Swedish statistics.

We utilize techniques that are well established in the natural sciences for measuring the relationship or difference of two vectors. We consider two statistics that measure the strength of the relationship between the different releases of the data: the Spearman and the Kendall Tau rank correlation coefficients. The relationship involves the different estimates of GDP that are available in the quarter following the one to which they refer. We also utilize two different measures of distance: Euclidean and Mahalanobis. They differ in the assumptions made about the statistical independence of the vectors. The Euclidean measure is only applicable to vectors that are independent.

1. Measures of Association

The two statistics (Spearman and Kendall's Tau) that calculate the strength of the relationship between the two vectors each measure different aspects of the data.[7] Both are concerned with the rankings of the growth rates of the components of the two vectors relative to each other. However, the former merely considers the two sets of ranks while the latter also shows whether the components of the vector have a tendency to move together. In both cases, the rankings are based on the growth rates of each component in each vector (vintage of the data).

a. Spearman Rank Correlation Coefficient

The Spearman Rank Correlation Coefficient is a non-parametric measure of the strength of a link between two sets of data, in this case the vintages of the first and later BEA estimates.

[7] An alternative measure, the Wilcoxon signed rank test, is inappropriate in our case because it assumes that the differences between the different vintages for each of the components are drawn from the same population. This is not an appropriate assumption due to the different variables used. For example, the scale is different for each component. We could also expect that the measurement errors from one vintage to another could differ depending on the component.

The reported growth rates of each component are converted into ranks and the differences, d_i, of each observation in the two vintages is calculated. The value of the Spearman coefficient (ρ) is then obtained from:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)}, \tag{4}$$

where n is the number of GDP components included in the vector and d_i is the difference between the ranks of the ith element of each vector.

b. Kendall Tau Rank Correlation Coefficient

The Spearman statistic only considers the difference in ranks that is assigned to each component within each vector, while Kendall's tau measures the tendency of two variables to change in the same direction.[8] Suppose that the first estimates of (the growth rate of) two GDP components are X_{1i} and X_{1j} and that the final (third) estimates of the same components are X_{3i} and X_{3j}. We rank each element in the vectors X_1 and X_3 by their size. Then, if the rank of X_{1i} minus the rank of X_{1j} has the same sign as the rank of X_{3i} minus the rank of X_{3j}, the pair of observations are said to be concordant. If the signs are opposite, the observations are called discordant. If there are 10 components of GDP in each vintage vector, then there are 45 pairs of such observations.[9] If the number of concordant (discordant) pairs is large (small), it would indicate a strong (weak) relationship between the two vintages of data.

The strength of relationship can be measured by the number of discordant pairs, which is reported as the Kendall tau distance. But this measure has a disadvantage because its size depends on the number of elements in the vector. Consequently, there is also a normalized tau

[8] Noether (1981) argues that the Kendall statistic is preferred because it is a better measure of the "strength of the relationship."
[9] If there are n observations, there would be $n(n-1)/2$ pairs.

distance, which is weighted by the number of pairs. This measure will thus run between 0 and 1, where 0 means the vectors have exactly the same ranking and 1 means that there is no agreement in the ranking. The Kendall tau (τ) rank correlation coefficient is also normalized by the number of pairs but runs between -1 and 1. A tau coefficient of 1 means that there is perfect agreement between the two rankings, -1 means that the rankings are exactly reversed, and 0 means they are independent. We will report the Kendall tau rank correlation coefficient here (the distance measures provide similar results). The formula is:

$$\tau = \frac{C - D}{\frac{n(n-1)}{2}}, \tag{5}$$

where C is the number of concordant pairs, D is the number of discordant pairs, and n is the number of components in each vector.

2. Distance Measures[10]

Assume that we have two independent vectors, X_1 and X_3, representing two vintages of data consisting of n components in each vector. The difference between the two vectors can be measured by the Euclidean distance between them:

$$d(X_1, X_3) = \sqrt{(X_1 - X_3)'(X_1 - X_3)}. \tag{6}$$

This procedure is only applicable to vectors that are independent. The vectors representing two vintages of data referring to the same quarter are (hopefully) not independent. If the vectors are not independent, we will use a different distance measure: Mahalanobis Distance.

When the two vectors are not independent, we use a distance measure which takes into account the relationship between the two vectors. This measure, that is a generalization of the

[10] See Abdi (2007) for a discussion of different distance measures.

Euclidean distance measure, takes into consideration the variance-covariance matrix of the data and is called the Mahalanobis distance.[11]

$$d_W^2(X_1, X_3) = (X_1 - X_3)' W (X_1 - X_3), \tag{7}$$

where W is the inverse of the sample variance-covariance matrix.

There is no intuitive economic interpretation of the distance measures that we obtain. Rather we will compare the magnitudes of these measures for periods when the economy was expanding with the numbers when the economy was in recession. For this calculation, we again follow Young (1987) and combine the data for the quarter before the peak and the quarter after the trough with the numbers for the quarters that constituted the recession as defined by the NBER. A difference of means test is then used to determine whether the differences were statistically significant.

For the Mahalanobis Distance we create separate weights for the recession periods versus the expansion periods for the difference in means test. These separate weights are based on the possibility that data for the recessions and expansions have different variance-covariance matrices. We, therefore, construct two separate weighting matrices; one is based on the sample variance-covariance of the observations recorded during expansions and the other is based on the sample variance-covariance of the observations recorded during recessions.

II. Data

We analyze the US nominal and real GNP/GDP data for the period 1970Q1-2010Q3. Since we are concerned with the compositional accuracy of the various vintages of the data, we examine both the headline GNP/GDP estimates and the ten components of GNP/GDP that have

[11] Mahanalobis distance is also associated with discriminant analysis. For other economic forecast applications of this measure, see Banternghansa and McCracken (2009) and Jordá et al (2010).

been published in real time in the *Survey of Current Business* throughout this period. These variables are: (1) durable consumption expenditures, (2) nondurable consumption expenditures, (3) personal services consumption expenditures, (4) nonresidential fixed investment, (5) residential fixed investment, (6) changes in business inventories, (7) exports, (8) imports, (9) federal government purchases, and (10) state and local government purchases.

Our real-time data were obtained from the ArchivaL Federal Reserve Economic Data (ALFRED®), maintained by the Federal Reserve Bank of St. Louis.[12] For both the nominal and real data, we create two vectors of these 10 series. They are constructed as the compound annual rate of change for each series, except for changes in business inventories which are measured in billions of dollars. The first vector represents the first release by the Bureau of Economic Analysis (BEA) of the estimates (prepared in the first month after the end of the quarter). The second vector represents the data available at the end of the third month after the end of the quarter. We will use the current BEA terminology for these different vintages by calling the first vector the "advance estimates" and the second vector the "third estimates."[13]

---

[12] To complete our dataset we supplemented what was available on ALFRED® with additional information available from the BEA's Survey of Current Business from the BEA's website for 1994 – 2010 (http://www.bea.gov/scb/date_guide.asp) and archived online back to 1921 by the Federal Reserve Archival System for Economic Research (FRASER®) maintained by the Federal Reserve Bank of St. Louis (http://fraser.stlouisfed.org/publications/SCB/).

[13] The timing and terminology of vintages of data released by the BEA have evolved over time. The current terminology is "advance" for the estimate released approximately 30 days after the end of the quarter, "second" for the estimate released approximately 60 days after the end of the quarter, and "third" for the estimate released approximately 90 days after the end of the quarter was adopted with the comprehensive revision released in July 2009 (Seskin and Smith, 2009). Previously, both the terminology as well as the timing of the releases varied. From 1988 until 2009, the timing of the three releases was similar to the current schedule, but the terminology was "advance" then "preliminary" then "final." Until 1988, the three estimates were released after each quarter on a 15-day, 45-day, and 75-day cycle. They were referred to alternatively as the 15-day, 45-day, and 75-day releases or "preliminary," "first revision," and "second revision." Prior to 1974 there were only the first two releases which were referred to simply as "preliminary" for the 15-day release and "final" for the 45-day release. The estimates began to be released later in the month in 1988 in response to a change in the schedule for processing monthly merchandise trade forms (Young, 1993).

III. Results

A. Bias

1. Nominal GDP

The bias test for the headline nominal GDP estimates yielded mixed results. The constant in the Mincer-Zarnowitz regression was not significantly different from zero, but the slope did not equal one. The Wald test rejected the joint null indicating that the earliest nominal GDP estimates were biased estimates of the numbers available 90 days after the quarter to which they refer (Table 1; line 1). What is more important is *that information about the state of the economy was incorporated into the earliest estimates*. This result is observed in the first line of Table 2 because the estimated coefficient associated with the dummy was not significantly different from zero. The results were similar using the Holden-Peel test applied to the data revisions.

Turning now to the estimates of the components, if the state of the economy is not considered, the null hypothesis was rejected at the 10% level for all but three of the components: inventories and both federal and state and local government expenditures (Table 1). If the state of the economy is included in the analysis, the null is again rejected in seven of the ten cases. The inventory estimates were now biased, but non-durable consumption no longer appears biased. The coefficient of the dummy variable was significant in several of the equations, indicating that the state of the economy affected estimates of those variables.

When the SUR procedure was used to test for the bias of the entire vector of nominal components, the null was rejected with a p value of less than 0.001, regardless whether or not the state of the economy was modeled. This was not surprising given that the null had been rejected for such a large percentage of the components.

2. Real GDP

Similar to the results for nominal GDP, the Mincer-Zarnowitz equation shows that the first **real** GDP estimates were biased estimates of the third numbers (Table 3, line 1) but that the coefficient on the dummy was not significant (Table 4, line 1). We further find that the mean absolute size of the revisions is 0.6% which is 25% of the mean absolute change in real GDP. With respect to the components, when the state of the economy was not taken into account, the estimates of only three variables were biased at the 10% level: durable consumption, non-residential investment, and exports (Table 3). When the effect of the state of the economy was taken into account, the null was rejected for six components at the 10% level[14], and the coefficient on the dummy variable was significant in three cases, consumption services, inventories, and imports (Table 4). We conclude that information about the state of the economy was not incorporated into the estimates of these particular components. It is interesting that the government estimates were never found to be biased, showing that BEA had more accurate information about these variables. The SUR approach again rejected the null that the vector of estimates was unbiased. The results again were similar using the Holden-Peel test applied to the data revisions.

B. Compositional Changes

1. Relationship (Correlation) Between Advance and Final Vintages

For both the nominal and real data, both the Spearman and Tau coefficients show that the changes in the components of the two vintages of data are closely related. Although the correlations appear to be generally lower in some recessionary periods, again there is no obvious

---

[14] In the private sector, only the real consumption of non-durable goods residential investment were found to be unbiased once we included the recession dummy.

cyclical pattern in the magnitudes of the coefficients (Figures 1 and 2). Although the numbers are not tabulated here, some interesting results obtained from the Tau coefficient should be noted.[15]

The lowest value of tau for the nominal variables was only 0.38. The lowest value of tau for the real data was only 0.29. We reject the null hypothesis of independence of the two vectors at the 5% level for all except 3 cases for the nominal data (1976Q2, 1991Q2 and 1992Q3) and 6 cases for the real data (1970Q3, 1976Q2, 1984Q2, 1990Q1, 1991Q3, and 2007Q1). Tau measures the number of concordant pairs, which in our sample, is large. *Thus, our results indicate that there is a strong relationship between the two vintages of data.*

2. Mahalanobis Distance

Because the two vintages of data are so closely related, we do not present the results using Euclidean Distance.[16] The Mahalanobis distance measure produced interesting results. For both the nominal and real GDP data, the mean of this measure was larger in recessions than in expansionary periods (Table 5). However, this difference was only significant when different weights were used for the two groups of periods, but then the difference was significant beyond the 0.001 level. Figures 3A and 3B show that this measure spikes during these recessionary periods, but it should also be noted that this distance measure was significantly smaller during the mild recession of 2001. This result indicates that revisions between the first and third estimates are most pronounced in the neighborhood of recessions.

If the recession of 2007 is excluded, the Mahalanobis distance measure is smaller in the second half of our sample period. This indicates that the revisions between the first and third estimates are smaller during this period. There are two possible explanations for these smaller

[15] The tabulated results are available from the authors upon request.
[16] The results for the Euclidean distance analysis are available from the authors upon request.

changes. First, it is possible that the procedures for making the first estimate are better. Alternatively, less new information may be forthcoming in the 60 days after the first estimate has been released.

IV. Conclusions

We have found that the headline nominal and real GNP/GDP numbers are both biased. But this bias is not attributable to BEA failing to include information about the state of the economy in its initial estimates. A number of the early estimates of both the nominal and real GDP components were also biased. Some of the component estimates were affected by the failure to include information about the state of the economy.

Our analysis of the measure estimates and the correlations showed that the early data generally reflected the composition of the changes in GDP that was observed in the later data. Thus, under most circumstances, an analyst could use the early data to obtain a realistic picture of what had happened in the economy in the previous quarter. However, the differences in the composition of the vectors of the two vintages were larger during recessions than in expansions. Unfortunately, it is in those periods when accurate information is most vital for nowcasting and forecasting.

Table 1
Mincer-Zarnowitz Regressions: Nominal GDP and Components
1970I – 2010III (Newey-West Standard Errors in Parentheses)

	Constant	Slope	Wald Test Probability
Nominal GNP/GDP (growth)	-0.059 (0.121)	$1.056^{\dagger\dagger\dagger}$ (0.019)	<0.001
Nominal Consumption Durable Goods (growth)	-0.095 (0.162)	$1.037^{\dagger\dagger}$ (0.017)	0.097
Nominal Consumption Non-Durable Goods (growth)	0.333* (0.181)	0.986 (0.021)	0.079
Nominal Consumption Services (growth)	0.358** (0.163)	$0.948^{\dagger\dagger}$ (0.023)	0.074
Nominal Fixed Investment Nonresidential (growth)	0.806*** (0.307)	1.021 (0.020)	0.002
Nominal Fixed Investment Residential (growth)	0.242 (0.246)	$1.040^{\dagger\dagger\dagger}$ (0.015)	0.017
Nominal Private Inventories (change in)	-0.167 (0.842)	0.988 (0.024)	0.832
Nominal Exports (growth)	1.843*** (0.501)	1.056 (0.042)	<0.001
Nominal Imports (growth)	0.942 (0.595)	1.027 (0.025)	0.065
Nominal Government Spending Federal (growth)	0.057 (0.244)	0.999 (0.037)	0.971
Nominal Government Spending State and Local (growth)	0.204 (0.184)	0.991 (0.023)	0.289
SUR of 10 Components			<0.001

*, **, ***: statistically significantly different from zero at the 10%, 5%, and 1% level respectively
$\dagger$, $\dagger\dagger$, $\dagger\dagger\dagger$: statistically significantly different from one at the 10%, 5%, and 1% level respectively

Table 2
Modified Mincer-Zarnowitz Regressions: Nominal GDP and Components
1970I – 2010III (Newey-West Standard Errors in Parentheses)

	Constant	Slope	Recession Dummy	Wald Test Probability
Nominal GNP/GDP (growth)	-0.067	$1.057^{\dagger\dagger\dagger}$	0.017	<0.001
	(0.159)	(0.021)	(0.163)	
Nominal Consumption Durable Goods (growth)	0.093	$1.034^{\dagger}$	-0.592	0.012
	(0.211)	(0.019)	(0.423)	
Nominal Consumption Non-Durable Goods (growth)	0.386*	0.985	-0.171	0.138
	(0.188)	(0.020)	(0.218)	
Nominal Consumption Services (growth)	0.384**	$0.955^{\dagger\dagger}$	-0.289**	0.020
	(0.153)	(0.022)	(0.139)	
Nominal Fixed Investment Nonresidential (growth)	1.036***	1.008	-0.551	0.002
	(0.354)	(0.023)	(0.728)	
Nominal Fixed Investment Residential (growth)	0.081	$1.0435^{\dagger\dagger\dagger}$	0.487	0.027
	(0.342)	(0.016)	(0.517)	
Nominal Private Inventories (change in)	1.810	0.956	-5.515***	0.071
	(1.119)	(0.029)	(2.083)	
Nominal Exports (growth)	1.757***	1.057	0.268	<0.001
	(0.670)	(0.045)	(1.058)	
Nominal Imports (growth)	1.900**	1.013	-2.874**	0.011
	(0.740)	(0.031)	(1.170)	
Nominal Government Spending Federal (growth)	-0.138	0.995	0.804	0.575
	(0.314)	(0.037)	(0.627)	
Nominal Government Spending State and Local (growth)	0.229	0.992	-0.105	0.459
	(0.202)	(0.023)	(0.204)	
SUR of 10 Components				<0.001

*, **, ***: statistically significantly different from zero at the 10%, 5%, and 1% level respectively
†, ††, †††: statistically significantly different from one at the 10%, 5%, and 1% level respectively

Table 3
Mincer-Zarnowitz Regressions: Real GDP and Components
970I – 2010III (Newey-West Standard Errors in Parentheses)

	Constant	Slope	Wald Test Probability
Real GNP/GDP (growth)	0.069 (0.077)	1.045††† (0.016)	<0.001
Real Consumption Durable Goods (growth)	-0.043 (0.150)	1.036†† (0.016)	0.072
Real Consumption Non-Durable Goods (growth)	0.216 (0.148)	0.988 (0.035)	0.236
Real Consumption Services (growth)	0.169 (0.117)	0.937† (0.035)	0.197
Real Fixed Investment Nonresidential (growth)	1.011*** (0.324)	0.994 (0.025)	0.001
Real Fixed Investment Residential (growth)	0.302 (0.311)	1.037†† (0.018)	0.134
Real Private Inventories (change in)	-0.363 (0.779)	0.992 (0.025)	0.810
Real Exports (growth)	1.683*** (0.478)	0.964 (0.031)	0.002
Real Imports (growth)	0.946* (0.534)	1.007 (0.025)	0.118
Real Government Spending Federal (growth)	-0.298 (0.262)	0.976 (0.026)	0.444
Real Government Spending State and Local (growth)	0.204* (0.113)	0.961 (0.026)	0.180
SUR of 10 Components			<0.001

*, **, ***: statistically significantly different from zero at the 10%, 5%, and 1% level respectively
†, ††, †††: statistically significantly different from one at the 10%, 5%, and 1% level respectively

Table 4
Modified Mincer-Zarnowitz Regressions: Real GDP and Components
1970I – 2010III (Newey-West Standard Errors in Parentheses)

	Constant	Slope	Recession Dummy	Wald Test Probability
Real GNP/GDP (growth)	0.062	$1.046^{\dagger\dagger}$	0.015	<0.001
	(0.101)	(0.022)	(0.168)	
Real Consumption Durable Goods (growth)	0.132	$1.031^{\dagger}$	-0.535	0.011
	(0.191)	(0.018)	(0.393)	
Real Consumption Non-Durable Goods (growth)	0.366**	0.975	-0.429*	0.121
	(0.162)	(0.033)	(0.234)	
Real Consumption Services (growth)	0.343**	$0.912^{\dagger\dagger}$	-0.342***	0.011
	(0.132)	(0.035)	(0.110)	
Real Fixed Investment Nonresidential (growth)	1.143***	0.985	-0.340	0.003
	(0.356)	(0.023)	(0.678)	
Real Fixed Investment Residential (growth)	0.033	$1.044^{\dagger\dagger}$	0.894	0.155
	(0.359)	(0.020)	(0.654)	
Real Private Inventories (change in)	1.337	0.963	-4.743**	0.090
	(1.024)	(0.029)	(1.899)	
Real Exports (growth)	2.006***	0.952	-0.981	0.003
	(0.595)	(0.033)	(1.063)	
Real Imports (growth)	2.278***	0.966	-3.782***	<0.001
	(0.608)	(0.027)	(0.992)	
Real Government Spending Federal (growth)	-0.327	0.976	0.103	0.617
	(0.298)	(0.026)	(0.737)	
Real Government Spending State and Local (growth)	0.236	0.959	-0.103	0.335
	(0.146)	(0.027)	(0.179)	
SUR of 10 Components				<0.001

*, **, ***: statistically significantly different from zero at the 10%, 5%, and 1% level respectively
†, ††, †††: statistically significantly different from one at the 10%, 5%, and 1% level respectively

Table 5
Mahalanobis Distance
Difference in Means Tests

	Full Sample Weights		**Different Weights Recessions v. Expansions**	
	Nominal	**Real**	**Nominal**	**Real**
Difference (Recession Mean minus Expansion Mean)	0.268	0.140	2.944***	2.346***
t-statistic	1.302	0.669	5.258	6.359
p-value	0.193	0.503	<0.001	<0.001

***Significant at the 1% level

Figure 1A

Kendall Tau Rank Correlation Coefficient for the Nominal Variables

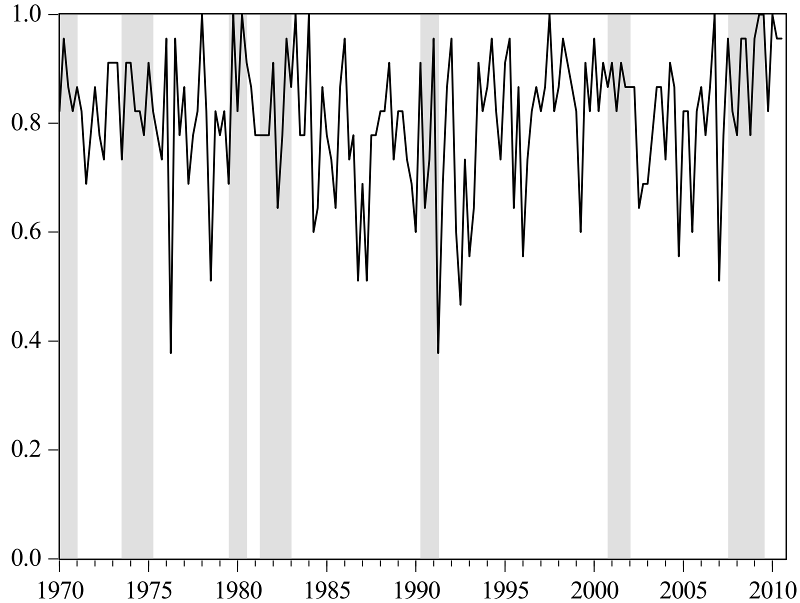
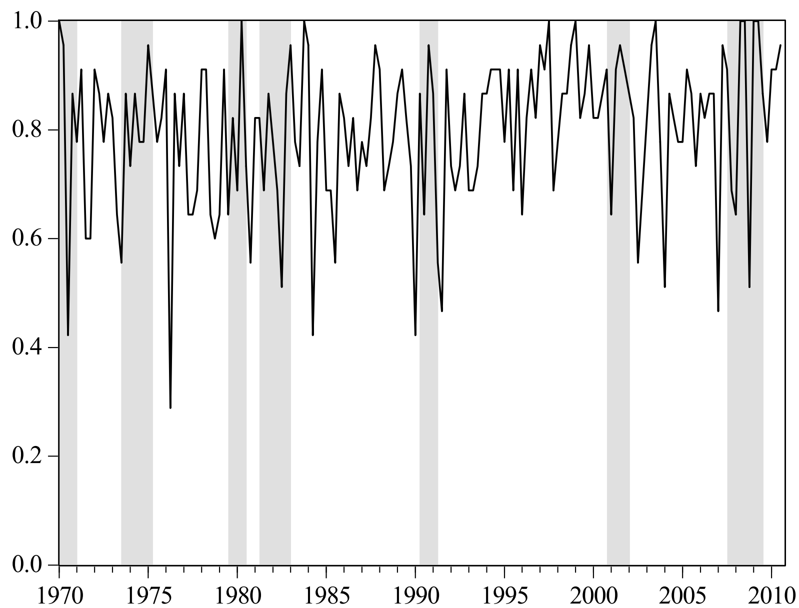


Figure 1B

Kendall Tau Rank Correlation Coefficient for the Real Variables



Note: Shaded bars in all graphs represent NBER dated recession plus one quarter on either side following Young (1993).

Figure 2A

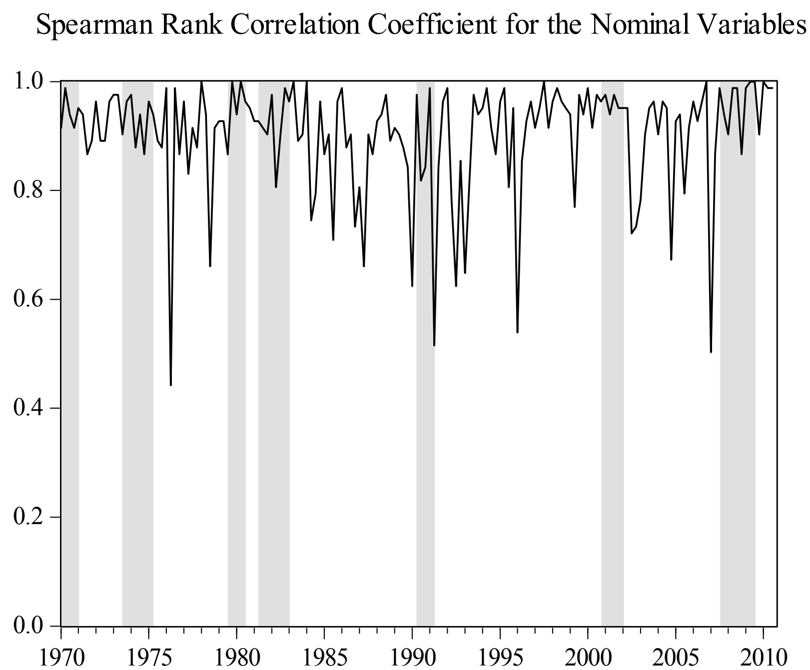


Figure 2B

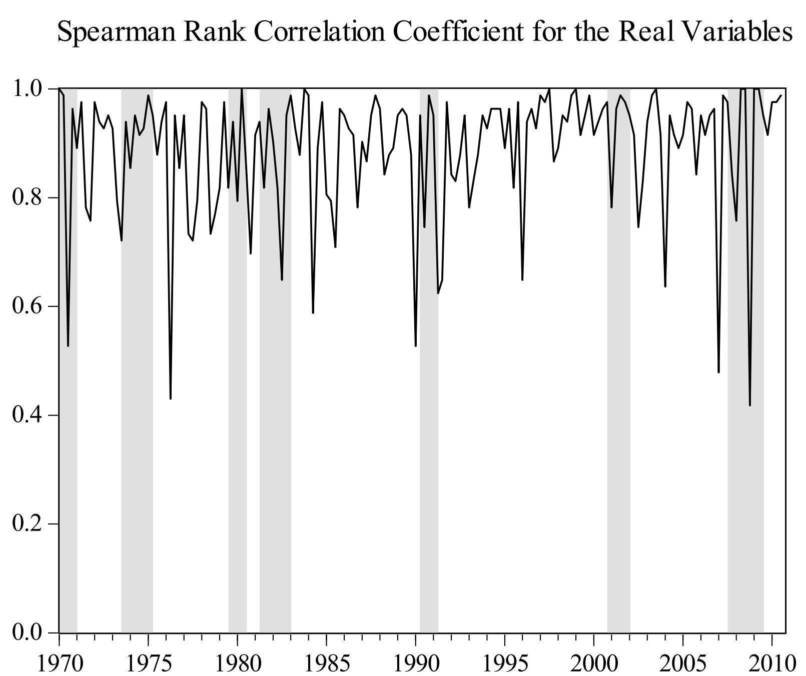
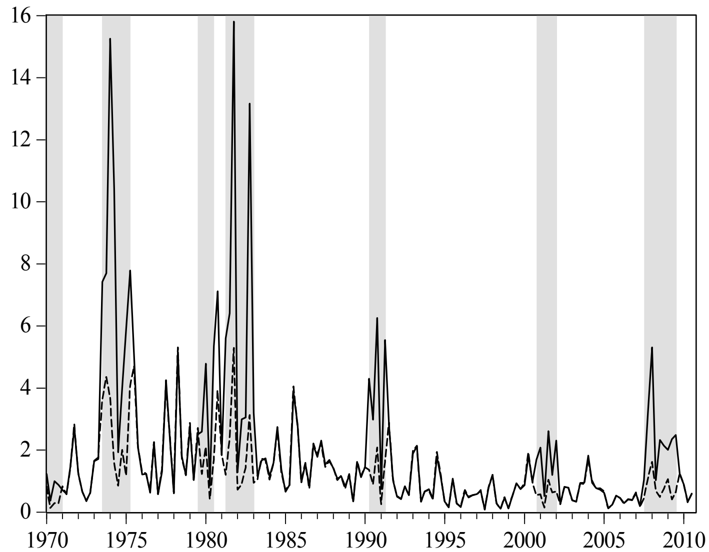


Figure 3A

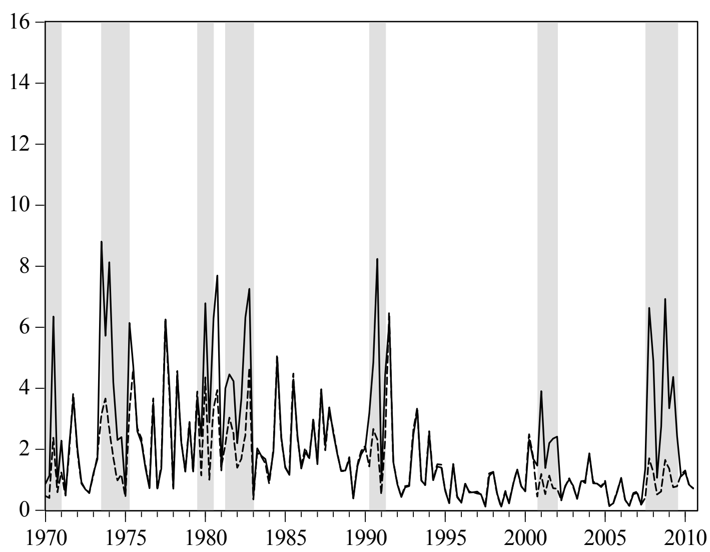
Mahalanobis Distance for the Nominal Variables



—— Mahalanobis Distance with Different Weights for Recessions versus Expansions
---- Mahalanobis Distance with Full Sample Weights

Figure 3B

Mahalanobis Distance for the Real Variables



—— Mahalanobis Distance with Different Weights for Recessions versus Expansions
---- Mahalanobis Distance with Full Sample Weights

References

Abdi, Herve (2007). Distance. In N.J. Salkind (Ed.) Encyclopedia of Measurement and Statistics. Thousand Oaks CA: Sage, pp. 280-284.

Aruoba, S. Borağan (2008). Data Revisions Are Not Well Behaved. Journal of Money, Credit, and Banking, 40, 319–340.

Banternghansa, Chanont and Michael W. McCracken (2009). Forecast Disagreement Among FOMC Members. Federal Reserve Bank of St. Louis Working Paper No. 2009-059.

Croushore, Dean, (2006). Forecasting with Real-Time Macroeconomic Data. In: Graham Elliott, Clive W. J. Granger, and Allan Timmermann, eds., Handbook of Economic Forecasting (Amsterdam: North-Holland), pp. 961–982.

Croushore, Dean, (2009). Frontiers of Real-Time Data Analysis. Forthcoming in Journal of Economic Literature.

Croushore, Dean, (2010). Forecasting with Real-Time Data Vintages. Forthcoming in: M. Clements and D. Hendry, eds., Oxford Handbook of Forecasting.

de Leeuw, Frank (1990). The Reliability of U.S. Gross National Product. Journal of Business and Economic Statistics, 8, 191-203.

Dynan, Karen E. and Douglas W. Elmendorf (2001). Do Provisional Estimates of Output Miss Economic Turning Points? Federal Reserve Board Finance and Economics Discussion Series No. 2001-52.

Faust, Jon, John H. Rogers and Jonathan H. Wright (2005). News and Noise in G-7 GDP Announcements. Journal of Money, Credit, and Banking, 37(3), 403-419.

Fixler, Dennis J. and Bruce T. Grimm (2002). Reliability of GDP and Related NIPA Estimates. Survey of Current Business, 82, January, 9-27.

Fixler, Dennis J. and Bruce T. Grimm (2005). Reliability of the NIPA Estimates of U.S. Economic Activity. Survey of Current Business, 85, February, 9-19.

Fixler, Dennis J. and Bruce T. Grimm (2008). The Reliability of the GDP and GDI Estimates. Survey of Current Business, 88, February, 16-32.

Grimm, Bruce T. and Robert P. Parker (1998). Reliability of the Quarterly and Annual Estimates of GDP and Gross Domestic Income. Survey of Current Business, 78, December, 12-21.

Groen, Jan J. J., George Kapetanios and Simon Price (2009). Real Time Evaluation of Bank of England Forecasts for Inflation and Growth. International Journal of Forecasting, 25, 74-80.

Hanson, M. S. and J. Whitehorn (2006). Reconsidering the Optimality of Federal Reserve Forecasts. Manuscript.

Holden, K. and D.A. Peel (1990). On testing for unbiasedness and efficiency of forecasts. Manchester School, 48, 120–127.

Joutz, Frederick L. and H. O. Stekler (1998). Data Revisions and Forecasting. Applied Economics, 30, August, 1011-1016.

Joutz, F. and H. Stekler (2000). An evaluation of the predictions of the Federal Reserve. International Journal of Forecasting, 16, 17-38.

Jordá, Òscar, Malte Knüppel, and Massimiliano Marcellino (2010). Empirical Simultaneous Confidence Regions for Path-Forecasts. European University Institute Economics Working Papers No. ECO2010/18.

Mankiw, N. Gregory, David E. Runkle and Matthew D. Shapiro (1984). Are Preliminary Announcements of the Money Stock Rational Forecasts? Journal of Monetary Economics, 14, 15-27.

Mankiw, N. Gregory and Matthew D. Shapiro (1986). News or Noise: An Analysis of GNP Revisions. Survey of Current Business, 66, May, 20-25.

McNees, Stephen K. (1986). Estimating GNP: The Trade-Off Between Timeliness and Accuracy. New England Economic Review, Jan./Feb., 3-10.

Mincer, J. and V. Zarnowitz (1969). The Evaluation of Economic Forecasts. In J. Mincer (Eds.) Economic Forecasts and Expectations. National Bureau of Economic Research, New York.

Morgenstern, Oskar (1963). On the Accuracy of Economic Observations, Princeton: Princeton University Press.

Mork, Knut A. (1987). Ain't Behavin': Forecast Errors and Measurement Errors in Early GNP Esimates. Journal of Business and Economic Statistics, 5, 165-175.

Neftci, S. N. and P. Theodossiou (1991). Properties and Stochastic Nature of B.E.A.'s Early Estimates of GNP. Journal of Economics and Business, 43, 231–239.

Noether, G. E. (1981). Why Kendall Tau? Teaching Statistics, 3, May, 41-43.

Öller, L.E. and A. Teterukovsky (2007). Quantifying the Quality of Macroeconomic Variables. International Journal of Forecasting, 23, 205-217.

Patterson, K. D. and S. M. Heravi (1991). The Information Content and Gain of Revisions to the Components of GDP. Applied Economics, 23, 90-912.

Patterson, K. D. and S. M. Heravi (2004). Revisions to Official Data on U.S. GNP: A Multivariate Assessment of Different Vintages. Journal of Official Statistics, 20, 573-602.

Sinclair, Tara M., Fred Joutz, and H. O. Stekler (2010). Can the Fed Predict the State of the Economy? Economics Letters, 108, 1, July, 28 – 32.

Seskin, Eugene P. and Shelly Smith (2009). Preview of the 2009 Comprehensive Revision of the National Income and Product Accounts: Changes in Definitions and Presentations. Survey of Current Business, 89, March, 10–27.

Swanson, N. R. and D. van Dijk (2006). Are Statistical Reporting Agencies Getting It Right? Data Rationality and Business Cycle Asymmetry, Journal of Business & Economic Statistics, 24(1), 24-42.

Stark, Tom and Dean Croushore (2002). Forecasting with a Real-Time Set for Macroeconomists. Journal of Macroeconomics, 24, 507-531.

Stekler, H. O. (1967). Data revisions and economic forecasting. Journal of the American Statistical Association**,** 62, 470–483.

Theil, Henri (1966). Applied Economic Forecasting, Amsterdam: North-Holland Publishing Company.

Young, Allan H. (1987). Evaluation of the GNP estimates. Survey of Current Business, 67, August, 18-42.

Young, Allan H. (1993). Reliability and Accuracy of the Quarterly Estimates of GDP. Survey of Current Business, 73, October, 29-43.

Zarnowitz, Victor (1982). On Functions, Quality, and Timeliness of Economic Information. Journal of Business, 55, 87–119.

Zellner, Arnold (1958). A Statistical Analysis of Provisional estimates of Gross national Product and its Components, of Selected National Income Components, and of Personal Saving. Journal of the American Statistical Association, 53, 54-65.