

JEREMY J. NALEWAIK

Board of Governors of the Federal Reserve System

The Income- and Expenditure-Side Estimates of U.S. Output Growth— An Update to 2011Q2

ABSTRACT In light of recent large revisions to the official measures of U.S. output, this update reviews the evidence in my 2010 Brookings Paper showing that the income-side estimate of output (currently called gross domestic income, or GDI) likely captures business cycle fluctuations in true output better than its better-known expenditure-side counterpart (called gross domestic product, or GDP). Most notably, over the 2007–09 downturn, the revisions moved the expenditure-side estimates closer to the income-side estimates, which showed that the downturn was considerably worse than reported initially by the expenditure-side estimates. The tendency for the expenditure-side estimates to be revised toward the income-side estimates is clearer now, as is a tendency for the smoothed income-side estimates to be revised away from the smoothed expenditure-side estimates.

Since publication of my 2010 Brookings Paper (Nalewaik 2010), the two official measures of U.S. output, GDP(E) and GDP(I), have passed through two annual revisions.¹ The revisions lowered GDP(E) by a sizable amount: GDP(E) in 2009Q3 was revised down about 2 percent relative to its level in early 2006. In light of these revisions, I update the evidence in Nalewaik (2010) showing that GDP(I) likely provides a more accurate estimate of output growth than GDP(E).

1. The official statistical agencies call the income-side estimate of output gross domestic income (GDI) and the expenditure-side estimate gross domestic product (GDP). Here, as in Nalewaik (2010), to make clear that GDI and GDP measure the same thing, I call the income-side estimate GDP(I) and the expenditure-side estimate GDP(E). Throughout the paper, “GDP” refers to real GDP.

Broadly speaking, three things have changed as a result of the revisions. First, some of the countercyclicality of the statistical discrepancy between GDP(E) and GDP(I), a striking finding in Nalewaik (2010), was revised away, but only with respect to the last few years. The countercyclicality from the mid-1980s to the mid-2000s remains.

Second, the tendency for the initial GDP(E) estimates to be revised toward the initial GDP(I) estimates has become clearer, strengthening the evidence that initial GDP(I) growth is a more accurate estimate of true output growth than is initial GDP(E) growth. In particular, GDP(I) picked up the onset and the severity of the 2007–09 downturn better and sooner than GDP(E), adding to the long list of cases where GDP(I) has recognized important economic phenomena before GDP(E). That list includes the productivity acceleration in the mid- to late 1990s and the sluggishness of the recovery following the 2001 recession.

Third, when the growth rates are smoothed into year-over-year changes, a tendency for the initial GDP(I) growth estimates to be revised away from the initial GDP(E) growth estimates has become clear as well. Indeed, the revisions to GDP(E) and GDP(I) growth tend to go in the same direction: if initial GDP(I) growth is above initial GDP(E) growth, both estimates tend to be revised up, and if initial GDP(I) growth is below initial GDP(E) growth, both tend to be revised down. This update discusses a potential explanation for this pattern in the revisions, related to differences in how well the estimates pick up output fluctuations from firm births and deaths.

Section I of this update discusses the changes to the estimates over the 2007–09 cyclical downturn since Nalewaik (2010). Section II discusses the implications of these changes for understanding the relative reliability of the latest available output growth rates. Section III discusses changes to the cyclicity of the statistical discrepancy. Section IV discusses the implications of these changes for understanding the relative reliability of the initial output growth rate estimates (those released about 3 months after the close of each quarter). Section V concludes with a brief discussion of progress on the recommendations of Nalewaik (2010) to improve the measurement of GDP. An online appendix interprets the variance of the estimates and the size of the revisions to the estimates, providing a response to assertions made in a recent article by Bureau of Economic Analysis (BEA) staff (Fixler, Greenaway-McGrevy, and Grimm 2011).²

2. Online appendixes for papers in this issue may be accessed at the Brookings Papers webpage, www.brookings.edu/economics/bpea.aspx, under “Past Editions.”

I. The Estimates over the 2007–09 Cyclical Downturn

The top panel of table 1 shows the revisions to the Q4-over-Q4 growth rates in GDP(E) and GDP(I) in 2007 and 2008 since Nalewaik (2010). Since that time, the initial (third) estimates of the Q4-over-Q4 growth rates in 2009 and 2010 have also become available and have passed through two and one annual revisions, respectively; these are shown in the bottom panel of table 1.

One remarkable feature of these recent revisions, discussed further in section IV, is the frequency with which they move Δ GDP(E) and Δ GDP(I) in the same direction. In particular, the downward revisions to the output growth estimates in 2007, 2008, and 2009 have continued and were particularly large for Δ GDP(E) in 2008. As can be seen in figure 1, which updates figure 6 in Nalewaik (2010), the peak in GDP(E) at that time was still 2 quarters late, in 2008Q2, but the current data now show that the peak in GDP(E) was 2007Q4, consistent with the NBER Business Cycle Dating Committee's call. These latest GDP(E) estimates show just how badly the initial GDP(E) estimates missed the onset and severity of the recession.

Based on the behavior of GDP(I), Nalewaik (2010) said that output over 2007–09 likely decelerated sooner and more sharply ahead of the

Table 1. Changes to Estimates of Growth in GDP(E) and GDP(I) since Nalewaik (2010)
Percent per year except where stated otherwise

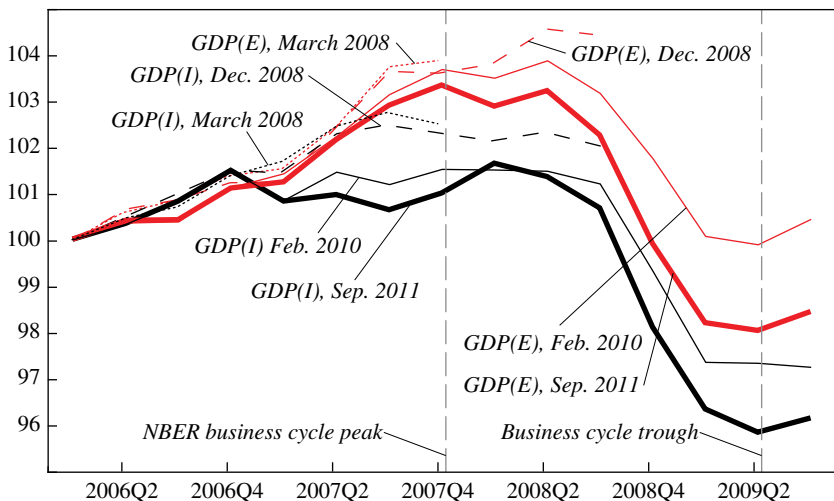
	<i>Estimate</i>		<i>Revision^a</i> <i>(percentage points)</i>
	<i>Spring 2010</i>	<i>Latest</i>	
<i>From 2006Q4 to 2007Q4</i>			
GDP(E) growth	2.5	2.2	–0.3
GDP(I) growth	0.1	–0.5	–0.5
<i>From 2007Q4 to 2008Q4</i>			
GDP(E) growth	–1.9	–3.3	–1.5
GDP(I) growth	–2.2	–2.9	–0.7
	<i>Initial (third)</i>	<i>Latest</i>	<i>Revision^a</i> <i>(percentage points)</i>
<i>From 2008Q4 to 2009Q4</i>			
GDP(E) growth	0.1	–0.5	–0.6
GDP(I) growth	–0.6	–0.9	–0.2
<i>From 2009Q4 to 2010Q4</i>			
GDP(E) growth	2.8	3.1	0.4
GDP(I) growth	3.1	3.5	0.4

Source: Bureau of Economic Analysis.

a. May differ from the difference between the previous two columns because of rounding.

Figure 1. GDP(E) and GDP(I) Estimates of Different Vintages, 2006Q1–2009Q3^a

Index, 2006Q1 = 100



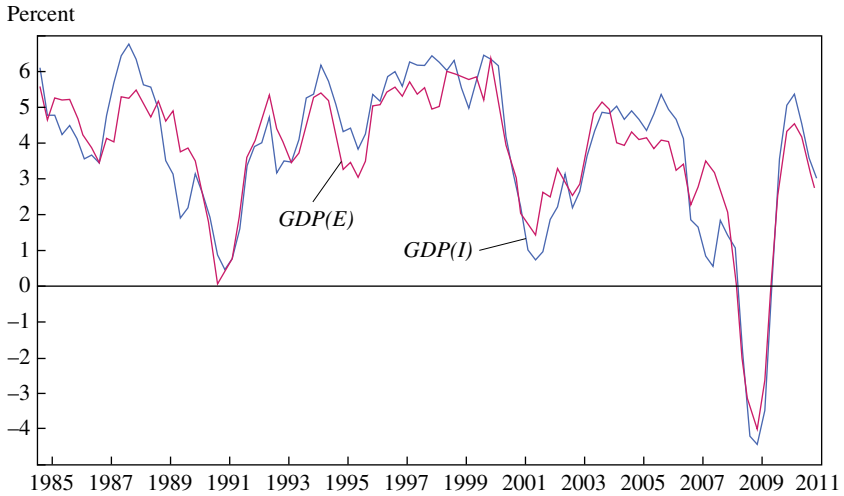
Source: Author's calculations using BEA data.

a. This figure updates figure 6 in Nalewaik (2010, p. 98).

recession, fell faster at the height of the recession, and recovered less quickly than was shown by GDP(E) at the time. The revisions have largely validated this claim: $\Delta\text{GDP}(E)$ was revised down toward $\Delta\text{GDP}(I)$ slightly in 2007 and more substantially in late 2008, at the height of the downturn.³ And whereas 2009Q3 $\Delta\text{GDP}(I)$ was revised from negative to positive, $\Delta\text{GDP}(E)$ was revised down in the second half of 2009, so that GDP(E) recovered less quickly than was initially estimated. These revisions to $\Delta\text{GDP}(E)$, both in the recession itself and in the early stages of the recovery, make the rise in the unemployment rate in 2009 less surprising than it seemed in early 2010.

Regarding the recovery, the news on output growth since Nalewaik (2010) has not been all bad. Initial $\Delta\text{GDP}(I)$ was above initial $\Delta\text{GDP}(E)$ in 2010, typically a good sign, as discussed in section IV. And some upward revisions to output growth have finally been reported. Figure 2, which updates figure 3 in Nalewaik (2010), shows the latest year-over-

3. At the time of Nalewaik (2010), average annualized $\Delta\text{GDP}(E)$ was -6 percent in 2008Q4 and 2009Q1, and $\Delta\text{GDP}(I)$ was about $-7\frac{1}{2}$ percent. Since then, average $\Delta\text{GDP}(E)$ has been revised down to about $-7\frac{3}{4}$ percent for those two quarters, and $\Delta\text{GDP}(I)$ has been revised down further as well, to about $-8\frac{1}{2}$ percent.

Figure 2. Year-over-Year Growth Rates of GDP(E) and GDP(I), 1985Q1–2011Q2^a

Source: Author's calculations using BEA data.

a. This figure updates figure 3 in Nalewaik (2010, p. 87). Calculations use latest available estimates of GDP(E) and GDP(I).

year growth rates, which reveal the income-side measure moving modestly above the expenditure-side measure starting in 2010Q1 and continuing through 2011Q2.

II. Information Content of the Latest Growth Rates

The recent revisions have not altered the fact that GDP(E) and GDP(I) paint widely divergent pictures of the behavior of the economy in 2007, and the sources of this discrepancy remain unclear. Perhaps capital gains are not being stripped out of GDP(I) properly.⁴ Nalewaik (2010) pointed out some problems with some of the components of GDP(E), suggesting that the component estimates were too high; although these estimates were

4. Taking capital gains out of the tax data used to compute corporate profits and proprietors' income seems like a relatively straightforward task, since tax forms have separate lines for capital gains, and companies typically have little incentive to misclassify capital gains as ordinary income, since tax treatments typically favor capital gains. The BEA has suggested, however, that the task might not be so straightforward. In any case, since capital gains are likely procyclical, overreporting of capital gains would reduce the cyclicity of GDP(I) relative to the cyclicity of true output (since measured capital gains are subtracted from output) and thus does not explain why GDP(I) is more cyclical than GDP(E).

Table 2. Correlations between and Variances of Initial and Latest Available Estimates of Growth in GDP(E) and GDP(I)^a

<i>Measure</i>	<i>Initial</i> $\Delta GDP(E)$	<i>Initial</i> $\Delta GDP(I)$	<i>Latest</i> $\Delta GDP(E)$	<i>Latest</i> $\Delta GDP(I)$
<i>Correlations, 1978Q3–2011Q2</i>				
Initial $\Delta GDP(E)$	1.00			
Initial $\Delta GDP(I)$	0.94	1.00		
Latest $\Delta GDP(E)$	0.84	0.82	1.00	
Latest $\Delta GDP(I)$	0.77	0.81	0.80	1.00
<i>Correlations, 1984Q3–2008Q4</i>				
Initial $\Delta GDP(E)$	1.00			
Initial $\Delta GDP(I)$	0.91	1.00		
Latest $\Delta GDP(E)$	0.76	0.72	1.00	
Latest $\Delta GDP(I)$	0.68	0.74	0.70	1.00
<i>Variances</i>				
	1978Q3–2011Q2		1984Q3–2008Q4	
Initial estimates	8.16	8.61	4.73	5.30
Latest estimates	9.73	10.54	6.11	7.39
Revision ^b (difference between latest and initial)	2.86	3.70	2.71	3.37

Source: Author's calculations using BEA data.

a. This table updates table 1 in Nalewaik (2010, p. 74). "Initial" estimates are those in the third BEA release for each quarter.

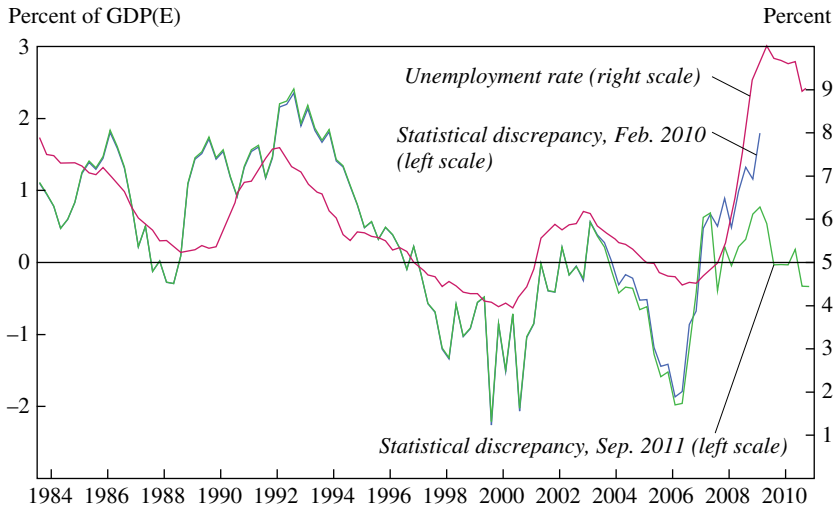
b. May differ from the difference between the previous two rows because of rounding.

later revised down, the revisions were not very large.⁵ For 2007, those components may remain badly mismeasured, as may much of personal consumption expenditures (PCE) on services.

Generalizing beyond 2007, B. Aruoba and others (2011) discuss optimal combinations of the latest GDP(E) and GDP(I) estimates. Nalewaik (2010) discussed how variances and revisions could be used to provide information on the relative reliability of the latest (not the initial) estimates. This analysis is little changed by the recent revisions: table 2, which updates table 1 of Nalewaik (2010), provides the new summary statistics, which

5. In particular, investment spending on residential improvements and personal consumption expenditures (PCE) on financial services were revised down by only \$4 billion and \$9 billion, respectively, not enough to be meaningful in explaining the 2007 annual change in the statistical discrepancy of about \$150 billion. However, the BEA seems to have agreed that there were problems in its methodology for estimating residential improvements, because the agency changed that methodology soon after Nalewaik (2010) was published. The BEA has also started folding in newly available Census data on financial services PCE, from the Service Annual Surveys and the Quarterly Services Surveys. However, since these new data begin in 2009, they are not helpful for determining what happened in 2007.

Figure 3. Statistical Discrepancy between GDP(E) and GDP(I) and the Unemployment Rate, 1984Q1–2011Q2^a



Source: Author’s calculations using BEA and Bureau of Labor Statistics data.

a. This figure updates figure 4 in Nalewaik (2010, p. 88). Estimates use latest available data as of September 2011.

are similar to those in Nalewaik (2010). Dennis Fixler, Ryan Greenaway-McGrevy, and Bruce Grimm (2011), in an article in the BEA’s *Survey of Current Business*, make a different set of assumptions about variances and revisions, and the online appendix to this update analyzes the validity of those assumptions.

III. The Cyclicity of the Statistical Discrepancy

Figure 3, which updates figure 4 in Nalewaik (2010), shows that the clear systematic relationship of the statistical discrepancy to the unemployment rate observed in that paper was largely revised away in 2008 and 2009. Whether the cyclicity of the discrepancy will appear again in the future may hinge on whether the lack of official annual surveys on much of the enormous services-producing sector of GDP(E) was the source of the cyclicity problem from the mid-1980s to the mid-2000s, and whether the availability of new annual services surveys starting in 2009 solves the problem. If that is not the case, the deviation from the pattern could be temporary, and indeed from 2009Q4 to 2011Q2 the pattern has

already emerged again with a level shift, as the statistical discrepancy and the unemployment rate have both drifted down. In any event, there remains a very large cyclical problem from the mid-1980s to the mid-2000s that the availability of new annual services data does not address; this problem may plague econometric analyses until it is fixed.

IV. The Information Content of the Initial Growth Rates

This section discusses the information content of the initial growth rates, with particular emphasis on the tendency for initial GDP(I) growth to predict revisions to initial GDP(E) growth.

IV.A. Quarterly Growth Rates

Table 3, which updates table 2 in Nalewaik (2010), extends the sample for most of the regressions out to 2011Q2, with little change to most of the original results that showed initial GDP(I) growth to be the better predictor of a wide variety of business cycle indicators, including the change in the unemployment rate in the current period and subsequent periods, employment growth (measured using the Current Population Survey) in current and subsequent periods,⁶ the manufacturing purchasing managers index in current and subsequent periods, changes in stock prices over previous periods, the slope of the Treasury yield curve in previous periods, and forecasts by the Survey of Professional Forecasters of GDP(E) growth itself from previous periods. Initial GDP(I) growth is also a better predictor of initial GDP(E) growth over the next quarter or two than is initial GDP(E) growth.

The results that have changed the most are the revision results using the short sample at the bottom of the table. Nalewaik (2010) used a sample ending in 2006Q4, ensuring that the revised estimates on the left-hand side of the regression incorporated all their major annual source data. Using the same rule, table 3 rolls the sample forward to end in 2008Q4. Using the 60 quarterly observations from 1994Q1 to 2008Q4, I find a statistically significant tendency for Δ GDP(E) to be revised toward initial Δ GDP(I). Indeed, when predicting latest quarterly Δ GDP(E) using initial Δ GDP(E) and initial Δ GDP(I), it has been optimal over this period to place almost 50 percent weight on GDP(I), up from about 30 to 40 percent in the sample employed in Nalewaik (2010).

6. Recall that there is no reason to suspect these labor market measures to be spuriously correlated with Δ GDP(I), since data from the CPS household survey are not used in the construction of Δ GDP(I).

Table 3. Regressions Testing the Predictive Content of Initial Estimates of Growth in GDP(E) and GDP(I), 1978Q3–2011Q2 Sample*

Dependent variable	Initial $\Delta GDP(E)$			Initial $\Delta GDP(I)$			Constant	Adjusted R ²
	t	t - 1	t - 2	t	t - 1	t - 2		
$\Delta GDP(E)_t$, latest estimate	0.73 (0.13)			0.19 (0.14)			0.23 (0.29)	0.71
	0.70 (0.12)	-0.22 (0.15)		0.16 (0.13)	0.32 (0.15)		0.11 (0.26)	0.72
		-0.45 (0.28)			0.92 (0.29)		1.31 (0.36)	0.24
$\Delta GDP(E)_t$, initial estimate		-0.30 (0.26)			0.72 (0.24)		1.41 (0.27)	0.21
			-0.23 (0.20)			0.41 (0.23)	2.06 (0.30)	0.04
$\Delta GDP(I)_t$, latest estimate	-0.00 (0.17)			0.90 (0.18)			0.29 (0.29)	0.65
	0.01 (0.18)	-0.03 (0.18)		0.84 (0.20)	0.12 (0.17)		0.16 (0.33)	0.65
		-0.11 (0.26)			0.62 (0.25)		1.29 (0.46)	0.21
$\Delta GDP(I)_t$, initial estimate		-0.09 (0.23)			0.58 (0.22)		1.34 (0.31)	0.24
			-0.28 (0.19)			0.51 (0.21)	2.04 (0.36)	0.06
$(UR_t - UR_{t-1}) \times 4$	0.00 (0.07)			-0.33 (0.08)			0.96 (0.17)	0.51
		0.04 (0.07)			-0.32 (0.08)		0.87 (0.21)	0.40
			0.00 (0.10)			-0.20 (0.12)	0.65 (0.23)	0.20

(continued)

Table 3. Regressions Testing the Predictive Content of Initial Estimates of Growth in GDP(E) and GDP(I), 1978Q3–2011Q2 Sample^a
(Continued)

Dependent variable	Initial $\Delta GDP(E)$			Initial $\Delta GDP(I)$			Constant	Adjusted R ²
	t	$t-1$	$t-2$	t	$t-1$	$t-2$		
$(E_{t-1}^{household}/E_{t-1}^{household})^4$	0.12 (0.11)	-0.05 (0.09)		0.30 (0.11)	0.43 (0.12)		0.05 (0.27)	0.40
			-0.17 (0.14)			0.41 (0.15)	0.09 (0.35)	0.38
ISM _{manuf.}	0.15 (0.39)	0.42 (0.37)		1.31 (0.40)	0.92 (0.36)		0.48 (0.41)	0.16
						0.60 (0.42)	47.61 (0.74)	0.49
							47.88 (0.71)	0.41
							49.42 (0.83)	0.13
$\text{Log}(S\&P500_t - S\&P500_{t-4})/4$	-0.13 (0.29)			0.78 (0.34)			0.22 (1.04)	0.19
$r_{t-8}^{Taxes(10)} - r_{t-8}^{Taxes(2)}$	-0.14 (0.06)			0.20 (0.07)			0.57 (0.16)	0.07

SPF forecast, $\Delta\text{GDP}(E)_{i,t}$	0.09 (0.11)	0.39 (0.11)	0.81 (0.28)	0.49
SPF forecast, $\Delta\text{GDP}(E)_{i,t-1}$	0.05 (0.06)	0.19 (0.08)	1.72 (0.29)	0.18
SPF forecast, $\Delta\text{GDP}(E)_{i,t-2}$	0.08 (0.06)	0.06 (0.08)	2.20 (0.32)	0.10
$\Delta\text{GDP}(E)$, latest, 1994Q1–2008Q4	0.53 ^b (0.14)	0.47 (0.14)	-0.39 (0.19)	
	0.54 (0.14)	0.47 (0.14)	-0.40 (0.39)	0.68
$\Delta\text{GDP}(I)$, latest, 1994Q1–2008Q4	-0.04 ^b (0.20)	1.04 (0.20)	-0.37 (0.26)	
	-0.04 (0.20)	1.04 (0.21)	-0.38 (0.36)	0.63

Source: Author's regressions.

a. This table updates table 2 in Nalewaik (2010, pp. 78–79). "Initial" estimates are those in the third BEA release for each quarter. Regressions are run on the full sample period unless stated otherwise. Numbers in parentheses are Newey-West standard errors using eight lags. UR = unemployment rate, $E^{\text{household}}$ = employment growth from the CPS household survey, ISM^{manufact} = purchasing managers' index from the Institute for Supply Management manufacturing survey, r^{treasury} = yield on n -year Treasury note, SPF = Survey of Professional Forecasters.

b. Regression coefficients are constrained to sum to 1.

IV.B. Q4-over-Q4 Growth Rates

When the quarterly growth rates are smoothed into Q4-over-Q4 growth rates, the evidence is even more favorable to GDP(I), as can be seen in figure 4, which updates figure 2 in Nalewaik (2010). The top panel shows that, since the mid-1990s, when predicting latest Q4-over-Q4 Δ GDP(E), it has been optimal to place 84 percent weight on initial Q4-over-Q4 Δ GDP(I) and only 16 percent weight on initial Q4-over-Q4 Δ GDP(E). The regression slope is highly significant (adjusting for limited degrees of freedom using a t distribution), 12 out of 17 data points are in the correct quadrants (northeast and southwest), and the adjusted R^2 is 0.37, up from 0.25 in Nalewaik (2010).

Interestingly, the bottom panel now shows a statistically significant tendency for Δ GDP(I) to be revised away from initial Δ GDP(E), with 12 out of 17 data points in the northwest and southeast quadrants and an adjusted R^2 of 0.26. If initial Δ GDP(I) is higher than initial Δ GDP(E), Δ GDP(I) tends to be revised up; with initial Δ GDP(E) tending to be revised toward initial Δ GDP(I), both estimates then tend to be revised up if initial Δ GDP(I) is higher than initial Δ GDP(E). Similarly, both estimates tend to be revised down if initial Δ GDP(I) is lower than initial Δ GDP(E). These results show that the gap between initial Δ GDP(I) and initial Δ GDP(E), averaged over several quarters, has been highly informative in recent years. A possible explanation is discussed below.⁷

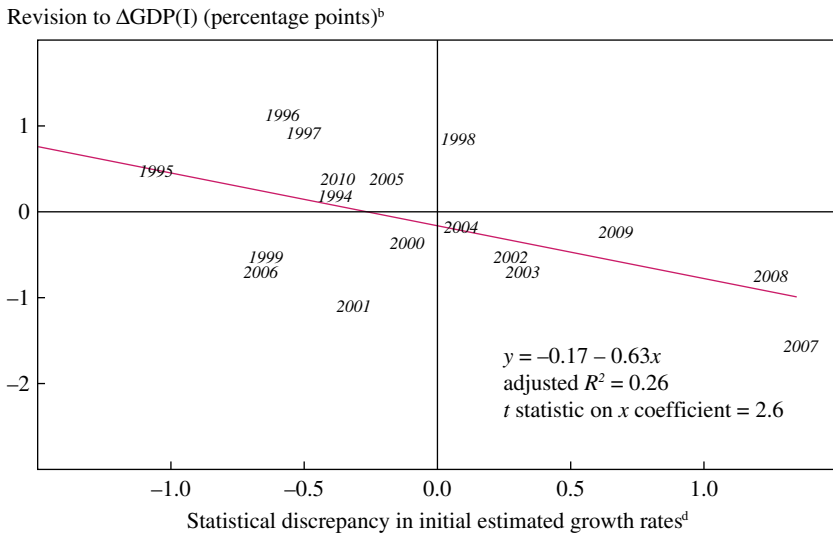
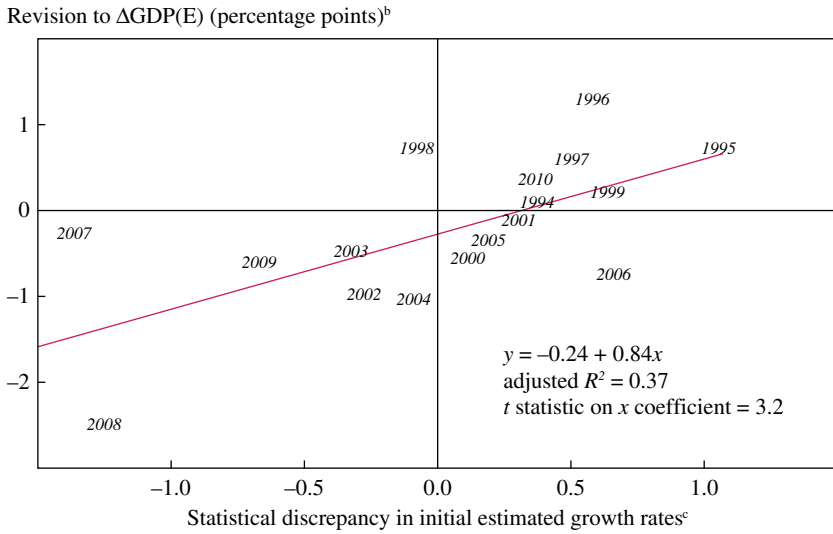
Naturally, the revisions to the two estimates have a positive correlation, over 0.6, and a rather striking 15 out of 17 revisions have been in the same direction for both estimates. The only exceptions are 1999 and 2005. But the magnitudes of the revisions can differ. Sometimes the estimates converge, as in 2008, when initial Δ GDP(E) and Δ GDP(I) started at around -1 percent and -2 percent, respectively, and then converged to around -3 percent, but typically the estimates do not converge but actually become more dissimilar, as in 2007.

IV.C. On the Gains to Placing Some Weight on Δ GDP(I)

The root-mean-square (RMS) of the Δ GDP(E) revisions plotted in figure 2 is around 0.9 and could be reduced to 0.7 using the regression results

7. These results counsel that if the BEA has in place any procedures for shrinking the initial Δ GDP(I) estimates toward the initial Δ GDP(E) estimates, those procedures should probably be discarded, or at least minimized so they have less of an effect on average growth rates over several quarters. That assumes that the BEA wants its initial Δ GDP(I) estimates to be optimal predictors of its latest, revised Δ GDP(I) estimates. If, in contrast, the BEA wants its initial Δ GDP(I) estimates to be optimal predictors of its latest, revised Δ GDP(E) estimates, the initial Δ GDP(I) estimates are fine the way they are.

Figure 4. Revisions to Output Growth Measures and the Statistical Discrepancy^a



Source: Author's calculations using BEA data.

a. This figure updates figure 2 in Nalewaik (2010, p. 80).

b. Difference between initial and latest available, fourth quarter to fourth quarter.

c. Difference between initial $\Delta\text{GDP(I)}$ and initial $\Delta\text{GDP(E)}$, fourth quarter to fourth quarter.

d. Difference between initial $\Delta\text{GDP(E)}$ and initial $\Delta\text{GDP(I)}$, fourth quarter to fourth quarter.

illustrated in figure 2.⁸ Is such a gain in the accuracy of the initial estimates worth the costs associated with changing current procedures? Couching the issue in terms of RMS errors does not do adequate justice to it, in my opinion, because initial $\Delta\text{GDP(I)}$ picked up some very important economic developments over the past couple of decades much better than initial $\Delta\text{GDP(E)}$. Given the BEA's historic emphasis on $\Delta\text{GDP(E)}$, modifying the initial $\Delta\text{GDP(E)}$ estimates would have allowed the economics profession and the broader public to recognize and reach consensus on these important developments sooner than they did in real time, which could have been quite valuable. The list of these important economic developments, all of which can be seen in figure 4, includes:

—*The onset and extent of the productivity acceleration in the mid- to late 1990s.* R. Anderson and K. Kliesen state that “the increasingly rapid productivity growth that began in the 1990s was the defining economic event of the decade” (Anderson and Kliesen 2010, p. 129). Although this characterization may be debatable, it was clearly a very important development. Some analysts did use GDP(I) to analyze productivity in real time (see the 1997 *Economic Report of the President*), but the issue was obscured by measurement error in $\Delta\text{GDP(E)}$. From 1994 to 1999, the initial Q4-over-Q4 GDP(E) growth rates were, on average, $\frac{1}{2}$ percentage point below the initial Q4-over-Q4 GDP(I) growth rates, and the Q4-over-Q4 GDP(E) growth rates ended up being revised up a little more than $\frac{1}{2}$ percentage point, on average, over this time period.

—*The sluggishness of output growth in the “jobless recovery” of late 2002 and early 2003.* Although this was apparent in the jobs data themselves in real time, it was less evident in initial $\Delta\text{GDP(E)}$ than in initial $\Delta\text{GDP(I)}$. Ultimately, much of the discordance between the output data and the jobs data (the subject of much discussion and analysis at the time) went away after subsequent downward revisions to both $\Delta\text{GDP(E)}$ and $\Delta\text{GDP(I)}$. In 2002 and 2003, both estimates were revised down by between $\frac{1}{2}$ and 1 percentage point per year.

—*The weakening of the economy ahead of the 2007–09 recession.* As discussed earlier, $\Delta\text{GDP(E)}$ was revised down toward $\Delta\text{GDP(I)}$ in 2007. Although the downward revisions are not particularly large, a best guess is

8. Similarly, the RMS of the $\Delta\text{GDP(I)}$ revisions is around 0.7, which could have been reduced to 0.6 by moving initial $\Delta\text{GDP(I)}$ away from initial $\Delta\text{GDP(E)}$. Note that although the variance of the revisions to quarterly $\Delta\text{GDP(I)}$ is higher than the variance of the revisions to quarterly $\Delta\text{GDP(E)}$ in table 1 (this is also true using a sample starting in 1994), the reverse is true using the Q4-over-Q4 growth rates.

Table 4. Contributions to the Latest Revision to the 2007Q4–2008Q4 Change in GDP(E)

Percentage points

<i>Component</i>	<i>Contribution^a</i>
PCE, durable goods	–0.2
PCE, nondurable goods	0.2
PCE, services	–0.7
Nonresidential structures	–0.3
Equipment and software	–0.3
Residential structures	–0.2
Change in private inventories	–0.6
Exports	–0.1
Imports	–0.3
Government	–0.1
Change in GDP(E)	–2.5

Source: Author's calculations from BEA data.

a. Components do not sum to the total change in GDP(E) because of rounding.

that at least part of the large remaining discrepancy between $\Delta\text{GDP(E)}$ and $\Delta\text{GDP(I)}$ in 2007 is due to unresolved measurement errors in $\Delta\text{GDP(E)}$.

—*The fact that the economy was in recession in the first half of 2008, and the severity of the downturn in late 2008.* This episode was discussed earlier.

It would seem worth a considerable cost to have official output data that pick up very important events such as these.

IV.D. A Closer Look at 2008

Before concluding, the enormous downward revisions to the initial 2008 $\Delta\text{GDP(E)}$ estimates are worthy of some additional attention. Table 4 breaks down the revision to 2008 Q4-over-Q4 $\Delta\text{GDP(E)}$:⁹ What is striking is the across-the-board nature of the downward revisions, with all major components revising down except nondurable goods PCE. What could have caused such large and widespread downward revisions? The potential problems with the initial estimates of $\Delta\text{GDP(E)}$ and $\Delta\text{GDP(I)}$ include missing data for some components, sampling errors, and non-sampling errors such as survey nonresponse. Sampling errors are random and should be uncorrelated across the range of surveys used to compute the different components, so this is an unlikely explanation. Missing data were likely part of the problem for services PCE, but as Steve Landefeld noted in his comment on Nalewaik (2010), the other components of GDP(E) are based on “some form of direct monthly or quarterly source data” (Landefeld 2010, p. 114).

9. The contributions are averages over the four quarters of 2008 of the quarterly contributions reported by the BEA.

So the widespread downward revisions are probably due to non-sampling errors correlated across components.

IV.E. One Potential Explanation: Firm Births and Deaths

One possible explanation, admittedly speculative, is the following: incomplete corrections for firm births and deaths, particularly of small businesses, have all the ingredients required to generate the correlated non-sampling errors that appeared in 2008. First, output fluctuations due to firm births and deaths are likely highly correlated with the aggregate business cycle, and through that channel, across components of GDP(E). Second, output fluctuations from firm births and deaths do not appear to be picked up by the monthly and quarterly surveys used to produce GDP(E). New firms are added with too long a lag, and when a firm drops out of the sampling frame, the survey administrator will not always know whether the firm has shut down (resulting in an imputed growth of -100 percent for that firm) or whether it has simply chosen to stop responding to the survey (which would require a different imputation), since the monthly and quarterly surveys used to produce GDP(E) are voluntary. The fraction of nonrespondents that are firm deaths likely varies with the business cycle, and although a birth-death model might pick up some of this, I know of no such model being applied to the expenditure surveys, so misclassification of nonrespondents is probably correlated with the business cycle. Finally, output fluctuations from firm births and deaths appear much more likely to be picked up by the annual surveys that the BEA folds into GDP(E) after the initial estimates. These annual surveys bring in new firms and are mandatory (although some nonresponse continues), allowing a better parsing of the fraction of firm nonrespondents that have shut down.

Although this is somewhat speculative as well, initial Δ GDP(I) may pick up firm births and deaths better than initial Δ GDP(E), for several reasons. First, in contrast to the major surveys used to compute GDP(E), the Current Employment Statistics (CES) survey actually has a birth-death model, facilitated by the availability of quarterly benchmarks from the Quarterly Census of Employment and Wages (QCEW), which should pick up firm births and deaths because it is a universe count of wage and salary payments derived from tax records.¹⁰ Second, the online appendix to Nalewaik (2010) shows that response rates to the CES are generally higher than those to the major surveys used to compute GDP(E), suggesting that misclassi-

10. The CES is the source for the initial estimates of wages and salaries and some components of initial GDP(E), but it is the source for a larger share of initial GDP(I), at least 40 percent.

fication of nonrespondents is less of a problem for the CES. Finally, since 2002, the initial Q4-over-Q4 growth rates of GDP(I) incorporate QCEW data into the first 3 quarters of the 4-quarter change, probably picking up a substantial amount of output variation due to firm births and deaths.

Finally, and most speculative of all, these conjectures may help explain patterns in the revisions to $\Delta\text{GDP(I)}$ and $\Delta\text{GDP(E)}$. Although initial $\Delta\text{GDP(I)}$ likely picks up some of the cyclical variation from firm births and deaths—and more than initial $\Delta\text{GDP(E)}$, which may pick up close to none—initial $\Delta\text{GDP(I)}$ probably does not pick up all of that cyclical variation. Then the tendency of $\Delta\text{GDP(I)}$ and $\Delta\text{GDP(E)}$ to revise in the same direction could be due to both of them incorporating cyclical variation from firm births and deaths. The revisions would tend to go in the direction of initial $\Delta\text{GDP(I)}$ minus $\Delta\text{GDP(E)}$, if part of that initial gap reflects cyclical variation from firm births and deaths picked up by $\Delta\text{GDP(I)}$ and missed by $\Delta\text{GDP(E)}$, variation that is likely positively correlated with the remaining cyclical variation from firm births and deaths that will appear later through revisions.

Of course, some other constellation of facts might explain the patterns in the revisions described here. Some additional analysis from the statistical agencies explaining these patterns would certainly be helpful.

V. Concluding Thoughts

Nalewaik (2010) concluded with some recommendations for improved GDP measurement, and I conclude here with an update on how action on those recommendations has progressed since then. The first suggestion was that the BEA report $\Delta\text{GDP(I)}$ in its press releases, preferably in table 1 along with $\Delta\text{GDP(E)}$, and provide some discussion of $\Delta\text{GDP(I)}$. Although the BEA has added real GDI growth (as the agency calls it) to its press release, it is currently reported in the very last table of the release, and the press release does not discuss it. More could be done on this front. A second recommendation was that the BEA feature as its headline output measure a 50–50 average of GDP(E) and GDP(I). The BEA has made it clear that it will not consider reporting such an average without a methodology for balancing the accounts, so that the components of GDP(E) and GDP(I) add up to the average.¹¹ This is a difficult task, but the BEA has begun

11. I argued in Nalewaik (2010) that this is not essential, because although $C + I + G + NX = \text{GDP(E)}$ now, the evidence strongly suggests that GDP(E) does not equal true GDP, so $C + I + G + NX \neq \text{GDP}$ now. Averaging without balancing would simply make that explicit. It is certainly understandable that the BEA would be concerned about nonadditivity, although it accepted nonadditivity in its real estimates when moving to chain weighting in the mid-1990s.

researching the issue and has suggested it might consider producing and publishing a set of averaged and balanced accounts as satellite accounts. A project like that would not be particularly costly—probably a lot less costly than some other projects the BEA is pursuing—and the evidence here and in Nalewaik (2010) suggests that it would be well worth the effort.

ACKNOWLEDGMENTS I thank Steve Braun, David Lebow, the editors, and participants at the Brookings Panel conference. The views expressed in this paper are solely those of the author and are not necessarily those of the Federal Reserve Board. The author reports no potential conflicts of interest.

References

- Anderson, R., and K. Kliesen. 2010. "FOMC Learning and Productivity Growth (1985–2003): A Reading of the Record." *Federal Reserve Bank of St. Louis Review* (March/April): 129–53.
- Aruoba, B., F. Diebold, J. Nalewaik, F. Schorfheide, and D. Song. 2011. "Improving GDP Measurement: A Forecast Combination Perspective." Working Paper no. 17421. Cambridge, Mass.: National Bureau of Economic Research.
- Fixler, D., R. Greenaway-McGrevy, and B. Grimm. 2011. "Revisions to GDP, GDI, and Their Major Components." *Survey of Current Business* (July): 9–31.
- Landefeld, J. Steven. 2010. "Comment [on 'The Income- and Expenditure-Side Estimates of Output Growth']." *BPEA* (Spring): 112–23.
- Nalewaik, Jeremy. 2010. "The Income- and Expenditure-Side Estimates of Output Growth." *BPEA* (Spring): 71–106.

Comment and Discussion

COMMENT BY

STEVEN BRAUN¹ Disputes over the relative merits of GDP(E) and GDP(I)—Jeremy Nalewaik’s nomenclature for real GDP and real GDI, respectively—as measures of aggregate economic activity have been largely settled. Although it commonly receives less attention than the headline GDP(E) measure, GDP(I) has independent value. Nalewaik’s 2010 Brookings Paper showed that initial estimates of real GDP(I) growth help predict revisions to real GDP(E) growth for the same period, whereas initial estimates of real GDP(E) growth are less helpful for predicting revisions to growth of real GDP(I). That paper also showed that GDP(I) is more highly correlated than GDP(E) with a variety of macroeconomic variables not directly related to the construction of either GDP(E) or GDP(I). The present update of Nalewaik (2010) shows that the two annual revisions and another year and a half of new data since the paper was written have reinforced those conclusions.

Much of this is not controversial, in the sense that Nalewaik and the staff of the Bureau of Economic Analysis (BEA) agree that both GDP(I) and GDP(E) have merit. In his comment on Nalewaik (2010), BEA Director Steven Landefeld noted that in making early estimates of both GDP(E) and GDP(I), BEA staff must fill in components for which hard data are missing with trended or judgmental estimates. One reason for preferring the product-side estimate is that it has fewer of these holes: 86 percent of the third estimate of GDP(E), but only 37 percent of the contemporaneous GDP(I) estimate, is based on some form of direct monthly or quarterly source data. One must, however, attach an important caveat to this characterization of

1. The views expressed are solely those of the author and not necessarily those of the Council of Economic Advisers or the Obama administration.

the two measures, namely, that the 86 percent figure counts data from the new Quarterly Services Survey (QSS) as part of the data available for estimating consumer spending on services (which constitutes 47 percent of GDP). Because most components of the QSS are still very new, most of the QSS time series have insufficient length to allow seasonal adjustment using standard methods. Until a longer time series has accumulated, for most parts of the QSS, the BEA must separate the trend from the seasonal fluctuations in consumer spending in a less satisfactory fashion. Even when the QSS is fully operational, it will still be important to verify the utility of the QSS for anticipating the Service Annual Survey, against which consumer spending on services is benchmarked.

Whatever the details, it is clear that more of GDP(E) rests on data-based estimates than is the case for GDP(I). And it is completely understandable why a statistical agency would wish to feature a data-based measure over one that rests more heavily on judgment. After all, the task of the statistical agencies includes occasionally reporting news that some fraction of the audience does not want to hear. At those times, having a GDP measure that is more directly based on data rather than on judgment has an obvious advantage. In retrospect, one can interpret Nalewaik's results concerning the accuracy of GDP(I) growth estimates, relative to GDP(E) estimates, as validating many of the judgments made by BEA staff to produce the income-side measure.

A recent paper by BEA staff economists Dennis Fixler, Ryan Greenaway-McGrevy, and Bruce Grimm (2011) notes that GDP(I) shows more variance in its quarterly growth rate and is more subject to revision than GDP(E). As Nalewaik points out, revisions that add information are not necessarily a bad thing. One can understand, however, that users of the data find large revisions disconcerting, and it is certainly reasonable to include the size of revisions as one among a number of criteria for evaluating a statistical measure.

Fixler and his coauthors (2011) go on to note that a pooled estimate that combines GDP(E) and GDP(I) has the desirable property of small revisions. In particular, their table 12 reports that a 67-33 weighted average of GDP(E) and GDP(I) has a lower mean absolute revision over the 27 years from 1983 to 2009 than any other linear combination of the two measures, and that between 1993 and 2009 a simple 50-50 average has the lowest mean absolute revision.

The BEA recently has elevated the public profile of GDP(I). Since July 2010, real GDP(I) growth has been published in appendix table A of the quarterly GDP press release, and beginning in November 2011, GDP(I) has

been discussed in a paragraph in the text of the release, for those releases that include an estimate of the latest quarter's GDP(I). (The BEA has always published timely estimates of real GDP(I) growth rates, but until recently these appeared in a location, National Income and Product Accounts table 1.7.1, likely to be consulted only by specialists.) There is some argument for going further and also publishing a weighted average of the two growth rates, a measure that may be the best estimate of current movements in the economy. According to Landefeld (2010), however, many users would like to see a fully balanced set of accounts, with expenditure- and income-side sums that match the pooled estimate. For GDP the expenditure- and income-side concepts match precisely, and so creating a pooled estimate is only a matter of averaging. But how would one apportion, for example, a strong growth rate of profits into the various expenditure-side components? Creating pooled estimates for the thousands of GDP(E) and GDP(I) components of the national income and product accounts would be difficult and undoubtedly would require some arbitrariness and considerable judgment. Alternatively, perhaps the public and the press could be educated to understand that the BEA reports two estimates of real GDP growth, that they differ because of measurement error, that the best estimate is likely to lie somewhere in the middle, and that the BEA has produced a pooled estimate, recommended for certain purposes, that lies in between the two competing measures, but only for the top-line GDP figure and not its components.

Nalewaik (2010) finds, and the present update confirms, that GDP(I) is better correlated with the unemployment rate than is GDP(E). One consequence is that using a pooled estimate of real GDP growth in the appropriate regression will create a lower standard error around an Okun's Law-based estimate of potential real GDP growth. To see this, consider a version of such a regression using annual data, where the fourth-quarter-to-fourth-quarter change in the unemployment rate (ΔUR) is regressed on the fourth-quarter-to-fourth-quarter percent change in real GDP ($\%y$):

$$\Delta UR = \alpha(\%y^* - \%y).$$

The estimation is nonlinear so that both of the interesting parameters, $\%y^*$ (potential real GDP growth) and α (the Okun's Law coefficient), can be estimated directly. When estimated over the 11 years 2000 to 2010, as in my table 1, a 50-50 weighted average of GDP(E) and GDP(I) produces a regression with a modestly higher R^2 , and consequently a smaller standard error around the estimate of potential GDP growth (although these differences are likely not statistically significant). This is relevant for macroeconomic forecasters because the growth rate of potential real

Table 1. Okun's Law Estimation of Potential GDP Growth^a

<i>GDP measure</i>	<i>Regression coefficient α</i>	<i>Growth in potential real GDP y^* (percent per year)</i>	<i>Standard error of y^*</i>	<i>R²</i>	<i>Standard error of the regression</i>
GDP(E) only	0.50	2.74	0.50	0.69	0.72
Simple average of GDP(E) and GDP(I)	0.51	2.68	0.44	0.75	0.65

Source: Author's regressions.

a. Results are from a nonlinear regression of the unemployment rate (fixed-weighted using 2006 weights for six major demographic groups: male and female, and ages 16 to 19, 20 to 24, and 25 and above) on the indicated GDP measure, using annual (Q4-to-Q4) data from 2000 to 2010.

GDP in the recent past is a good place to begin consideration of real GDP growth in the near future.

The productivity estimates maintained by the Bureau of Labor Statistics (BLS) also could be redesigned to rest on a pooled expenditure- and income-side estimate of output growth. This could produce a better real-time estimate of productivity. For example, an income-side measure of nonfarm business productivity showed the first real-time evidence of the post-1995 pickup in productivity growth as shown by a chart in the 1997 *Economic Report of the President* (p. 74). Because the major productivity measures refer to either the private business or the nonfarm business sector rather than the entire economy, this would require that either BEA or BLS produce a pooled estimate of the relevant expenditure- and income-side measures of output, but this is something that could easily be done.

Judging the current position of the economy is a prerequisite for forecasting where it is going, and more information can help in figuring that out. Forecasters are always looking out a small and dirty window to guess which way the wind is blowing. If another window is available—in this case GDP(I)—it could improve the forecast. Indeed, not just forecasters but all econometricians who use GDP in their regressions may find that a 50-50 average of the expenditure- and income-side estimates serves better than the headline GDP(E) measure, in the sense of producing better-fitting results.

REFERENCES FOR THE BRAUN COMMENT

- Fixler, D., R. Greenaway-McGrevy, and B. Grimm. 2011. "Revisions to GDP, GDI, and their Major Components." *Survey of Current Business* (July): 9–31.
- Landefeld, J. Steven. 2010. "Comment [on 'On the Income- and Expenditure-Side Measures of Output']." *BPEA*, no. 1: 112–23.
- Nalewaik, Jeremy. 2010. "On the Income- and Expenditure-Side Measures of Output." *BPEA*, no. 1: 71–106.

GENERAL DISCUSSION Steven Landefeld began the discussion by presenting the Bureau of Economic Analysis' perspective on the issues raised by Nalewaik's research. Although the agency did not disagree with the practice of constructing weighted averages of GDP(E) and GDP(I)—GDP and GDI in official parlance—it viewed the potential gains as modest, reducing the mean absolute revision in the early GDP estimates from 1.3 percentage points to 1.1 or 1.0. A practical concern, corroborated by discussions both with users and with the agency's Australian counterpart, with publishing an average as the headline number was the difficulty of tracking back from the published number to the underlying monthly indicators.

Landefeld thought Nalewaik had summarized well the divergence between the two measures, particularly in 2007–08, but that the general contours of the recent recession were the same whichever measure, and whichever revision of that measure, one used. All measures showed that recession to have been the deepest, and the recovery from it the slowest, in the postwar period, and the revision in the cumulative decline in GDP(E) was only from –4 to –5 percent, although at least one quarterly revision, that for 2008Q4, was considerably larger. Landefeld presented a chart showing that the divergence Nalewaik had noted in the 2007–08 data had actually become somewhat larger in the most recently revised data.

Landefeld also reported on some changes the BEA was making to improve its estimates of GDP(E). Much of the problem with the GDP(E) data, the agency believed, lay with some of the monthly and quarterly indicators used to produce the early GDP estimates, rather than with the benchmarks, which come from the Economic Census, or the mandatory surveys. In particular, the BEA was investigating to what extent the observed downward pattern of revisions could be explained by nonresponse bias, and specifically by fewer less well performing firms than better-performing firms responding to the Census Bureau's nonmandatory monthly surveys. The agency was also working to incorporate data from the new Quarterly Services Survey: in the two revisions since the Spring 2010 meeting at which Nalewaik's original paper was presented, the BEA has replaced trend and indicator data for services—which are used to estimate nearly a quarter of consumer spending and one-fifth of GDP—with quarterly expenditure data from the new Census survey.

With respect to GDP(I), the agency's concern, as Nalewaik had noted, was with the large variances and revisions. The BEA believed that much of this was due to too much cyclicity in the numbers, for example arising from the unintended inclusion of capital gains. Capital gains are a change in price rather than in output and therefore should be excluded from output

estimates, and because they vary with the business cycle, they distort the estimates more in expansions than in recessions. Other potential cyclical biases relate to the treatment of intangibles, stock options, transfer pricing by multinational corporations, and income shifting for tax reasons.

Landefeld concluded by saying that the BEA's ability to improve its statistical reporting was hampered by the threat of budget cuts, especially to those programs like the Economic Census that are essential statistical infrastructure for the nation.

Following up on Steven Braun's comparison of GDP(E)- and GDP(I)-based estimates of productivity growth, Robert Gordon reported his own estimates using a measure he called "unconventional productivity," which has GDP(I) rather than GDP(E) in the numerator and household-reported hours worked instead of payroll hours in the denominator. Gordon argued that using this measure largely solves the puzzle of why productivity growth seemed to continue at a rapid pace even after the dot-com boom collapsed. Whereas the conventional measure showed average annual productivity growth of 2.1 percent during both 1996–2001 and 2001–04, his unconventional measure showed it falling from 2.3 percent in the first period to 1.3 percent in the second—a 50 percent drop. Even a measure using a simple average of GDP(E) and GDP(I) in the numerator showed a substantial difference. In other words, the productivity puzzle largely disappears when GDP(I) is used. Although Gordon refrained from concluding that the BEA should report an average, he did think that GDP(I) deserved greater prominence in the BEA reports than it currently receives.

John Quiggin quipped that there were only three things wrong with GDP as an output measure: it was gross, it was domestic, and it measured product. To begin, if one wants to measure productivity, then a net rather than a gross output measure is called for—it makes no sense to measure productivity with a statistic that includes depreciation replacement. Similarly, when one is analyzing various other macroeconomic phenomena, a measure of national rather than domestic output is more appropriate. GDP is useful when one is thinking about Okun's Law, but that hardly amounts to a rationale for publishing it as a headline number. Quiggin further claimed that the choice of measure was by no means merely an academic question—he had seen cases where the wrong choice led to policy errors, for example to the unwarranted conclusion that capital should be taxed less.

Justin Wolfers sought to clarify Nalewaik's point that the greater variance in the GDP(I) data was not an argument against using that measure: that would be like saying that when two economists disagree, one should always believe the one who revises her findings less than the other.

Wolfers saw the choice of the preferred output measure as reducible to three options: follow current practice and stick with GDP(E), use an average of GDP(E) and GDP(I), or abandon GDP(E) and switch to using GDP(I) only. He noted that Nalewaik's results, taken as a whole, clearly reject the first option and mildly reject the second, but that not a single one of Nalewaik's regressions falsified the hypothesis that one should use GDP(I) alone. Recent history, Wolfers argued, supported this conclusion: the focus on GDP(E) had left U.S. policymakers as late as September 2008 wondering whether the economy was in recession, as the most recent available data then showed GDP(E) still rising, but not GDP(I). Thus, the stakes in making the right choice of output measure are high.

Finally, Wolfers noted that Steven Landefeld had said at the Spring 2010 Panel conference where Nalewaik's original paper was presented that the BEA planned to do more to highlight GDP(I) in its official publications. But Wolfers was still hearing from reasonably sophisticated users of economic data who were having trouble finding the GDP(I) numbers in the releases. This suggested that more remained to be done.

David Romer corroborated Wolfers's observation that even otherwise knowledgeable users were still having trouble finding the GDP(I) numbers or were even unaware that they were published. He doubted that any user would object to immediately giving the GDP(I) numbers more prominence, specifically by publishing them in the lead table of the quarterly release. In the long run, he hoped, the BEA would go considerably further.

Stephanie Aaronson suggested that if the users of macroeconomic data were to put more focus on the GDP(I) numbers in their own work, it would both foster public understanding of the concept and encourage the statistical agencies to devote more resources to their construction and dissemination. Meanwhile she was already seeing some private forecasters use GDP(I) in their productivity models with great success.

Phillip Swagel confirmed Wolfers's claim that the reliance on GDP(E) had contributed to some confusion in senior policy circles about the state of the economy in late 2007. At the Treasury, Secretary Henry Paulson became persuaded that the economy was in worse shape than the numbers were indicating, from information that included his discussions with top corporate executives over the December holidays. Fortunately, in their deliberations senior policymakers do not rely only on GDP(E) but also gather a wealth of other information, both formal and anecdotal.

Responding to the discussion, Jeremy Nalewaik said he agreed with Landefeld that there was also measurement error in the income-side data:

for example, the profits numbers might be contaminated to some degree by capital gains. For that reason, even though some of his regressions indicated putting a 100 percent weight on GDP(I), he himself did not propose going that far.

Nalewaik did question Landefeld's contention that the differences between the numbers were small: for example, his regression of final on revised numbers using fourth-quarter-over-fourth-quarter growth rates had an R^2 of around 0.4, indicating that use of GDP(I) would reduce the size of subsequent revisions by about 40 percent.

Finally, Nalewaik made the point that variance in time-series estimates was not inherently bad. Variance in GDP estimates that stems from classical measurement error is clearly undesirable. If, however, one were to estimate GDP growth in a way that systematically varied less than true GDP growth, it would imply that one's estimates suffered from some non-classical kind of measurement error. He argued that the initial GDP(E) estimates in 2008 suffered from such error. Nonclassical measurement error that masks real variance in true GDP growth could be more dangerous than classical measurement error, which simply adds noise, since it could lull policymakers into a false sense of security. Additionally, non-classical measurement error is less well understood by economists than the classical kind.