

The Income- and Expenditure-Side Estimates of U.S. Output Growth

Jeremy J. Nalewaik*

April 15, 2010

Abstract

The U.S. produces two conceptually identical official measures of its economic output, currently called Gross Domestic Product (GDP) and Gross Domestic Income (GDI). These two measures have shown markedly different business cycle fluctuations over the past twenty-five years, with GDI showing a more-pronounced cycle than GDP. The goal of this paper is to determine which measure better reflects the business cycle fluctuations in true output growth, and a broad range of results favor GDI. GDI currently shows the 2007-2009 downturn was considerably worse than is reflected in GDP.

JEL classification: C1, C82.

Keywords: GDP, statistical discrepancy, news and noise, signal-to-noise ratios, business cycles

*Economist, Federal Reserve Board, 20th Street and Constitution Avenue, Washington, DC 20551 (e-mail: jeremy.j.nalewaik@frb.gov). Thanks to the economists at the U.S. Bureau of Economic Analysis, who answered many, many questions as I worked on this paper. Thanks for helpful comments to Andrew Figura, Dennis Fixler, Bruce Grimm, Steve Landefeld, David Lebow, Deb Lindner, Jonathan Millar, Raven Molloy, Brent Moulton, David Romer, Dan Sichel, Stacey Tevlin, Justin Wolfers, and participants at the March 2010 Brookings Panel on Economic Activity. Jessica Chan provided able research assistance. The views expressed in this paper are solely those of the authors and are not necessarily those of the Federal Reserve Board.

1 Introduction

The U.S. produces two conceptually identical official measures of its economic output, currently called Gross Domestic Product (GDP) and Gross Domestic Income (GDI). These two measures have shown markedly different business cycle fluctuations over the past twenty-five years, with GDI showing a more-pronounced cycle than GDP. These differences have become particularly glaring over the latest cyclical downturn, which appears considerably worse along several dimensions when looking at GDI. The aim of this paper is to determine which measure better represents the actual business cyclical fluctuations in output growth, and a wide variety of results suggest the answer is GDI.

In discussing the information content of these two sets of estimates, the confusion often starts with the nomenclature. GDP can mean either the true output variable of interest, or an estimate of that output variable based on the expenditure approach. Since these are two very different things, using “GDP” for both is confusing. Furthermore, since GDI has a different name than GDP, it may not be initially clear that GDI measures the same concept as GDP, using the equally valid income approach. So, to keep things straight, this paper refers to the true variable of interest as true output, the expenditure-side estimate of true output as GDP(E), and the income-side estimate of true output as GDP(I).

The paper presents results for both the initial output growth estimates available in real time, and the later estimates that have passed through more revisions. After presenting some basic facts about the estimates, section 3 discusses the initial growth rates and shows numerous results favoring GDP(I) growth. First, there is some evidence that the initial GDP(I) growth predicts revisions to GDP(E) growth, and no tendency for GDP(E) growth to predict revisions to GDP(I) growth. Second, initial GDP(I) growth is the better predictor of a wide variety of business cycle indicators that should be correlated with true output growth. These include all measures of output growth in subsequent periods, the change in the unemployment rate in the current period and subsequent periods, employment growth

(measured using a household survey) in current and subsequent periods, 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 of GDP(E) growth itself from previous periods. Each of these results suggest GDP(E) growth is either the noisier measure of true output growth or misses fluctuations in true output growth that appear in both GDP(I) growth and the other business cycle indicators. Third, initial GDP(I) growth has identified the onset of the last few cyclical downturns more quickly than initial GDP(E).

Section 4 discusses the latest revised growth rates. The section first establishes some basic facts about the discrepancies between the fully-revised estimates. On average, GDP(I) tends to grow faster than GDP(E) when the economy is expanding robustly, and GDP(I) growth falls below GDP(E) growth in recessions and in periods where the economy is sluggish. As such, the statistical discrepancy is highly negatively correlated with the business cycle. Why is this the case? A thorough analysis of the nature of the source data suggests that GDP(E) misses part of the business cycle, and that GDP(I) captures the business cycle better. Statistical analyses reach the same conclusion. First, the nature of the revisions suggests they add cyclical variation to GDP(I) that is not added to GDP(E), implying GDP(E) misses some cyclical variation. And second, the latest GDP(I) growth estimates are more highly correlated with a wide range of business cycle indicators, including changes in unemployment, the growth rate of employment, purchasing manager surveys (both manufacturing and non-manufacturing), changes in stock prices over previous periods, the slope of the treasury yield curve in previous periods, the high-yield bond spread from previous periods, and indicator variables for NBER recessions.

Section 5 discusses the behavior of the estimates over the most recent cyclical downturn. Output decelerated sooner, fell at a faster rate at the height of the downturn, and recovered less quickly when measured by GDP(I). Drawing on the results from the previous sections

and online Appendices, this section discusses how GDP(E) may have missed the severity of the downturn. Section 6 concludes with thoughts about the implications of the results in the paper for both data users and the BEA.

2 Basic Facts about the Estimates

The BEA’s first GDP(E) estimate for the most recent quarter, called the “advance” estimate, is released about a month after the quarter closes. Estimates of most components of GDP(I) for that quarter are included in the “advance” release, but the BEA is not comfortable releasing estimates of corporate profits or net income from the rest of the world at that time. GDP(I) first appears with the “second” release about two months after the quarter ends, except the estimates for fourth quarters, when GDP(I) first appears with the “third” release about three months after the quarter ends. To work with a complete time series of the initial growth rates, we focus on these “third” release estimates. However, in an online appendix, I repeat the regression results in section 3 using the “second” estimates for quarters where GDP(I) is available, and “advance” GDP(I) estimates constructed using the available income-side components and forecasts of corporate profits and net income from the rest of the world.

After the BEA releases its initial estimates of GDP(E) and GDP(I) growth for any given quarter, the estimates are revised numerous times. Table 1 shows the variances and correlations of the initial “third” estimates and the latest estimates which have passed through more revisions; here and throughout the paper $\Delta GDP(E)$ and $\Delta GDP(I)$ are short for the annualized quarterly growth rates of the estimates. We focus on two samples here. The first starts in 1978Q3, and is dictated by the start date of the time series of “third” growth rates employed in the paper, which is based on a real time dataset constructed by the BEA starting in 1978. When analyzing the latest, revised estimates, the paper

focuses on a shorter sample starting in the mid-1980s, because the divergences between the estimates are particularly stark and highly cyclical over this time period.¹ This second sample ends in 2006Q4 to ensure that the latest estimates have revised to fully incorporate all their major annual source data. The data from the mid-1980s to present are plotted in Figures 1 and 2.

The top two panels of table 1 show that the correlation of the initial estimates with the latest estimates is fairly high (0.85 and 0.82 for one sample, and 0.68 and 0.66 for the other). Nonetheless, the revisions do change the estimates in important ways. First, the bottom panel shows the variance of the revisions is somewhat larger for $\Delta GDP(I)$ than for $\Delta GDP(E)$. Moreover, the revisions tend to increase the variance of $\Delta GDP(I)$ more than the variance of $\Delta GDP(E)$. This suggests that the revisions add information to latest $\Delta GDP(I)$ that is not added to latest $\Delta GDP(E)$. Finally, and perhaps counterintuitively, the revisions tend to make the two measures less similar, reducing their correlation from 0.90 to 0.60 in the shorter sample.

Given the important differences between the latest estimates after they have passed through their revisions, this paper investigates two questions. The first is: what is the relative information content of the initial growth rates of GDP(E) and GDI(I)? Put differently, how much weight should we place on each of these initial growth rates? The second question is: what is the relative information content of the GDP(E) and GDI(I) growth rates after they have passed through all their revisions? In other words, how much weight should we place on each of these latest, revised growth rates?

Appendix A provides more background information about GDP(E) and GDP(I). Appendix B discusses the source data used to construct the initial growth rates, while Appendix C describes the source data incorporated at annual and benchmark revisions.

¹The precise start date chosen here is the econometric break point marking the beginning of the once widely-accepted phenomenon known as the Great Moderation. The precise start date is not particularly important, however; any start date around the mid-1980s gives similar results for the latest estimates.

3 The Information Content of the Initial Growth Rates

A detailed examination of the source data used to compute the initial “third” growth rates of $\Delta GDP(E)$ and $\Delta GDP(I)$ shows that both estimates suffer from similar types of measurement-error problems; see Appendix B. These problems include missing data for a substantial portion of each estimate, sampling errors, and non-sampling errors such as incomplete coverage, survey non-response, and incomplete corrections for firm births and deaths. A compelling case for the superiority of either estimate cannot be made based on such a detailed examination of source data, so this section dives right into more-informative statistical tests.²

Table 2 reports the main regression results examining the information content of the initial “third” growth rates. A good initial place to start is an examination of the predictive power of the initial estimates for the latest, revised estimates incorporating superior source data. Over the full sample, the initial $\Delta GDP(E)$ estimates predict well the latest estimates of $\Delta GDP(E)$, with the initial $\Delta GDP(I)$ adding little after conditioning on initial $\Delta GDP(E)$.³ Similarly, initial $\Delta GDP(I)$ predicts latest $\Delta GDP(I)$, with initial $\Delta GDP(E)$ adding little information after conditioning on initial $\Delta GDP(I)$. However, the final two blocs of regressions in table 2 show results for a sample starting in 1994Q1; we stop this subsample in 2006Q4 to ensure the latest estimates have passed through all their annual revisions, but extending the subsample to 2008Q4 produces similar results. The

²In his comments, J. Steven Landefeld suggests that, for the “third” estimates, a much greater fraction of GDP(I) than GDP(E) is based on “judgmental trends” instead of early source data. Almost all of the source data used to compute the “third” estimates is flawed and unrepresentative in some way, and breaking down the data using such a binary classification scheme is a highly subjective exercise. The detailed discussion of the source data in Appendix B suggests the evidence is less favorable to GDP(E) than this classification scheme suggests. Moreover, if a much greater fraction of GDP(I) were based on trends, we should expect “third” $\Delta GDP(I)$ to be much less variable than “third” $\Delta GDP(E)$, because trends should have less variance than the actual source data. The summary statistics in table 1 show that this is not the case; “third” $\Delta GDP(I)$ is actually slightly more variable than “third” $\Delta GDP(E)$.

³Newey-West standard errors using eight lags are in parentheses below the estimates.

first specification constrains the coefficients on initial $\Delta GDP(E)$ and initial $\Delta GDP(I)$ to sum to one, while the second does not; the results show that when initial $\Delta GDP(I)$ is one percentage point above initial $\Delta GDP(E)$, initial $\Delta GDP(E)$ has revised up about a third to two-fifths of a percentage point, on average, over this time period. Fixler and Grimm (2006) also find some tendency for initial $\Delta GDP(E)$ to revise towards initial $\Delta GDP(I)$, using a broader set of conditioning variables. The last bloc of results shows that, over this subsample, there remains no significant tendency for initial $\Delta GDP(E)$ to predict latest $\Delta GDP(I)$.

The initial $\Delta GDP(I)$ may have predicted revisions to the $\Delta GDP(E)$ over this sample because they are less noisy than the initial $\Delta GDP(E)$ estimates, or because they contain information about true output growth missed by the initial $\Delta GDP(E)$ estimates but incorporated into latest $\Delta GDP(E)$ through revisions. Both of these explanations are likely part of the story. Averaging the data into year-over-year growth rates eliminates much of the noise in the quarterly data and shows the plausibility of the second explanation. These year-over-year growth rates for 4th quarters are plotted in Figure 3 (this picture was first suggested to me by Bill Wascher). Broadly speaking, two periods drive the positive relation plotted here.⁴

First, during the mid-to-late 1990s, the gap between initial $\Delta GDP(I)$ and $\Delta GDP(E)$ was consistently positive, as the initial estimates showed GDP(I) growing faster than GDP(E). This phenomenon was discussed in real time at the Council of Economic Advisors (see the Economic Report of the President, 1997, pp. 72-74), the Federal Reserve Board (see Greenspan, 2004), and at the BEA itself (see Moulton, 2000), with conclusions generally favorable to GDP(I). Those conclusions were vindicated, since, ultimately, $\Delta GDP(E)$

⁴The line plotted is the predicted values from regressing the 13 Q4-over-Q4 growth rates of real GDP(E) on a constant and the gap between the initial (third) estimates of Q4-over-Q4 GDP(I) and GDP(E) growth. The coefficient on the gap is 0.98, with a standard error of 0.44, and an adjusted R^2 of 0.25. We also experimented with corrections that removed the effects of major methodological changes from the revisions, which increased the R^2 .

revised up towards initial $\Delta GDP(I)$. Initial $\Delta GDP(I)$ accurately captured information about the brisk pace of economic growth that was missed by the initial $\Delta GDP(E)$ estimates and incorporated only later through revisions (and probably only partially—see section 4). Second, in the period after the 2001 recession, in 2002 and 2003, the initial estimates of $\Delta GDP(I)$ showed a more sluggish recovery than the initial estimates of $\Delta GDP(E)$, so the gap between the initial estimates was negative.⁵ Ultimately, $\Delta GDP(E)$ revised towards initial $\Delta GDP(I)$ again; the recovery was indeed quite sluggish, and this information was reflected in $\Delta GDP(I)$ before it appeared in $\Delta GDP(E)$. Figure 4 shows no tendency for $\Delta GDP(I)$ to revise towards $\Delta GDP(E)$; if anything, $\Delta GDP(I)$ tends to revise in the opposite direction of the initial gap between $\Delta GDP(E)$ and $\Delta GDP(I)$ over this time period.

This particular set of revision results occurs over a short sample, and should be taken with a grain of salt. However, as a robustness check, we used past issues of the Survey of Current Business to extend back our sample, and results reported in online Appendix D show a marginally statistically significant tendency for initial $\Delta GDP(E)$ to revise towards initial $\Delta GDP(I)$ over this long sample extending from 1966Q4 to 2009Q3. And, after the data have passed through their first annual revision, a statistically significant tendency for initial $\Delta GDP(E)$ to revise towards initial $\Delta GDP(I)$ in subsequent revisions appears yet again, using the 1978Q3 to 2009Q3 sample. So it would probably be unwise to ignore these revision results entirely.

A decision to ignore the revision results implies that the weight that data users should place on the initial estimates is entirely determined by the weight placed on the latest, fully-revised estimates. An analyst who believes that latest $\Delta GDP(E)$ is more accurate

⁵As Appendix C outlines, since 2002, the BEA has incorporated information from the Quarterly Census of Employment and Wages (QCEW) into their wage and salary estimates a couple of months after their “third” estimate. These QCEW revisions provided much of the information to $\Delta GDP(I)$ on the relative sluggishness of the recovery; see the discussion in Nalewaik (2007a). The year-over-year growth rates for 4th quarters available in real time reflect these QCEW revisions.

than latest $\Delta GDP(I)$ should believe that initial $\Delta GDP(E)$ is more accurate than initial $\Delta GDP(I)$, and vice versa. So the results outlined in the next section, addressing the paper’s second question, are also critical to answering the first.

However, we can make considerable further progress on the paper’s first question directly, by examining the predictive power of the initial estimates for other important cyclical indicators. Broadly speaking, these regressions help establish which estimate is more informative about the business cycle, but the regressions also help answer the narrower question of which is the better estimate of true output growth. The inferior estimate of true output growth, containing relatively more noise or classical measurement error, should have a lower signal-to-noise ratio and be the inferior predictor of cyclical indicators correlated with true output growth, all else equal. This assumes the noise in the output growth estimates is uncorrelated with the measurement error in the cyclical indicators, and I have chosen cyclical indicators carefully to avoid this problem. An estimate may be inferior not only because it is noisier, but also because it misses more fluctuations in true output growth—contains less news or signal about true output growth—than the other estimate. But, again, such an inferior estimate should have a lower signal-to-noise ratio and be the inferior predictor of cyclical indicators that reflect those missing fluctuations in true output growth.

Returning to the top of table 2, we see that as a cyclical indicator of where output growth is headed, initial $\Delta GDP(I)$ is superior to initial $\Delta GDP(E)$. The initial estimates of $\Delta GDP(I)$ are positively related to output growth in the next quarter, whether output growth next quarter is measured by $\Delta GDP(E)$ or $\Delta GDP(I)$, initial or latest. Conditional on initial $\Delta GDP(I)$, initial $\Delta GDP(E)$ contains no information about output growth next quarter, and may actually be negatively related to output growth next quarter. This result holds two quarters ahead as well, when output growth is measured by either initial estimate. Following the logic outlined above, these results may obtain because initial $\Delta GDP(E)$ is noisier than initial $\Delta GDP(I)$, obscuring its signal about true output growth in subsequent

periods, or because initial $\Delta GDP(E)$ misses some of the shocks that produce serially-correlated fluctuations in true output growth, shocks that appear in initial $\Delta GDP(I)$.

We examine next the relation of the initial estimates to other cyclical variables that should be correlated with true output growth. These other variables should not be used in the construction of either $GDP(E)$ or $GDP(I)$, to avoid correlated measurement errors and spurious correlation. As outlined in appendices, the $GDP(E)$ and $GDP(I)$ estimates make little use of the Current Population Survey (CPS), a monthly household survey used to produce the unemployment rate.⁶ As one of the most important indicators of the business cycle, the unemployment rate is a good variable to use as a starting point for this analysis.

Table 2 shows that contemporaneously, initial $\Delta GDP(I)$ has a strong negative relation with the change in the unemployment rate, and negatively predicts changes to the unemployment rate one and two quarters ahead, while the coefficients on initial $\Delta GDP(E)$ are insignificant and have the wrong sign when conditioning on initial $\Delta GDP(I)$. Again, this may be because initial $\Delta GDP(E)$ is noisier than initial $\Delta GDP(I)$, or because initial $\Delta GDP(E)$ misses fluctuations in true output that both appear in $\Delta GDP(I)$ and are reflected in the differenced unemployment rate.

The next bloc of regressions show results using quarterly annualized employment growth computed from the household survey data, adjusted for breaks introduced by Census updates to the population. Initial $\Delta GDP(I)$ is positively related to employment growth this quarter, as well as one- and two-quarters ahead, while initial $\Delta GDP(E)$ contains little additional information about employment growth beyond that contained in initial $\Delta GDP(I)$.

⁶At first blush, some analysts might suspect that $GDP(I)$ must be more correlated with the unemployment rate than $GDP(E)$, because “income” is in the name $GDP(I)$ and the unemployment rate is a labor market concept. However, this reasoning is incorrect. Of the various components of the two output measures, we may expect based on *a priori* considerations that compensation will have higher-than-average correlation with unemployment, but the other components of $GDP(I)$ should then have lower-than-average correlation, since all the components of $GDP(I)$ add up to the same conceptual measure of output as $GDP(E)$. For example, stories in the press recently have suggested that some of the recent rebound in corporate profits was facilitated by weakness in the labor market, allowing firms to cut compensation costs.

Broadening out the results beyond labor market variables, the next bloc of regressions uses the Purchasing Managers' Index (PMI) from the Institute for Supply Management (ISM) manufacturing survey. The ISM measure is computed quite differently from $GDP(E)$ and $GDP(I)$; it is an aggregation of several diffusion indexes, so even though the companies participating in the ISM also participate in the surveys used to estimate $GDP(E)$ and $GDP(I)$, the measurement errors likely behave quite differently. Initial $\Delta GDP(I)$ explains the contemporaneous, one quarter ahead, and two quarters ahead movements in the ISM measure better than initial $\Delta GDP(E)$, with initial $\Delta GDP(E)$ providing no statistically significant information conditional on initial $\Delta GDP(I)$.

Business cycle analysts use a host of other variables to predict $\Delta GDP(E)$, most notably different asset prices, and since these asset prices are not used in the construction of the output growth estimates, they are prime candidate variables for testing the information content of the initial estimates. However, asset prices typically predict output growth in subsequent quarters, rather than being predicted by output growth, so to get the timing correct, we regress lagged values of these asset prices on the two initial output growth measures. This is a somewhat odd specification, but still quite instructive. The results essentially tell us which initial estimate is more consistent with market expectations of the business cycle from earlier periods.

The first asset-price specification regresses the log change in the S&P 500 stock price index from the end of quarter $t - 4$ to the end of quarter t , on the two initial output growth measures in quarter t . Initial $\Delta GDP(I)$ is strongly positively related to this current and lagged stock price change, while the coefficient on initial $\Delta GDP(E)$ is insignificant and negative. The next specification examines the slope of the yield curve, measured as the difference in yields between ten- and two-year treasury notes. This variable is most closely related to the output growth measures about two years hence; a regression of this measure from quarter $t - 8$ on the two initial output growth measures in quarter t yields a coefficient

on initial $\Delta GDP(I)$ that is significant with the correct (positive) sign, and a coefficient on initial $\Delta GDP(E)$ that is significant but the wrong sign.

The final set of testing variables employed here are median forecasts of output growth from the Survey of Professional Forecasters (SPF). These forecasters are trying to predict initial $\Delta GDP(E)$, presumably inclusive of any measurement errors in $\Delta GDP(E)$. However, if the forecasters do not yet have access to the source data used to compute the quarter of $\Delta GDP(E)$ they are trying to predict, their forecasts will likely reflect general information about the state of the economy which may be better related to initial $\Delta GDP(I)$ than to initial $\Delta GDP(E)$. This may be the case even for the current-quarter forecasts, because the survey occurs relatively early in the quarter before the analysts have much $GDP(E)$ source data. And the results show that those current quarter forecast are well explained by initial $\Delta GDP(I)$, with initial $\Delta GDP(E)$ providing no incremental explanatory power. The SPF forecasts for quarter t , made in the first half of quarter $t - 1$ are also better explained by initial $\Delta GDP(I)$ in period t than initial $\Delta GDP(E)$ in period t . Forecasters' expectations for how the economy will move in the current quarter and next quarter appear to play out more fully in the initial $\Delta GDP(I)$ estimates than in initial $\Delta GDP(E)$.

Given the tighter relation of initial $\Delta GDP(I)$ to all these business cycle indicators, a business cycle analyst placing full weight on the initial $\Delta GDP(E)$ estimates and no weight on the initial $\Delta GDP(I)$ estimates must: (1) care about true output growth in the current quarter only (i.e. not care about the business cycle more broadly or even where true output growth is headed next quarter), (2) believe the latest $\Delta GDP(E)$ estimates reflect all available information about true output growth, so neither latest $\Delta GDP(I)$ nor any other variable provides any additional marginal information about true output growth, (3) believe the superior explanatory power of initial $\Delta GDP(I)$ for all these other cyclical indicators tells us nothing about the relative accuracy of initial $\Delta GDP(I)$ and initial $\Delta GDP(E)$ as estimates of true output growth, and (4) discount entirely the revisions evidence. This

third point is clearly a stretch, and could only be the case if initial $\Delta GDP(I)$ contained variation uncorrelated to true output growth but correlated with all the other dependent variables employed in table 2, including actual forecasts of output growth. A much more plausible explanation is that initial $\Delta GDP(I)$ is more highly correlated with true output growth than initial $\Delta GDP(E)$, and that true output growth is correlated with all these other cyclical indicators. The second point above is a stretch as well, and would only be the case if all other variables, including latest $\Delta GDP(I)$, provided no information about true output growth beyond that contained in latest $\Delta GDP(E)$. That is quite an extreme position in favor of the accuracy of latest $\Delta GDP(E)$, and the results in the next section suggest latest $\Delta GDP(I)$ does contain a considerable amount information about true output growth missed by latest $\Delta GDP(E)$. Regarding the first point, this may be a reasonable position for the BEA to take. For analysts, true output growth may the only variable of interest for some purposes, but for other purposes this will not be the case.

The regression results in table 2 are broadly consistent with those in Nalewaik (2007a), who uses Markov switching models to show that $\Delta GDP(I)$ identifies cyclical turning points more quickly than $\Delta GDP(E)$ in real time. Specifically, at the NBER-defined start of the 1980, 1981-2, 1990-1 and 2001 recessions, real-time estimates of a Markov switching model using $\Delta GDP(E)$ alone put the odds that the economy was in a low-growth state at 52%, 40%, 45%, and 23%, respectively. Adding $\Delta GDP(I)$ to the model produced much-more accurate probabilities: 78%, 44%, 72%, and 70%. Most of the research in Nalewaik (2007a) was carried out in 2005, and the subsequent cyclical downturn was the first out-of-sample test of the main hypotheses of the paper. The model using $\Delta GDP(I)$ again performed much better around the start of the downturn in real time, and also performed better than some popular models using monthly indicators; see section 5.

While this section has focused on the “third” growth rates, the information content of the preceding “advance” and “second” growth rates are of critical importance for analysis

in real time, and Appendix D reports results for these vintages, as well as results for the estimates once they have passed through their first annual revision. Briefly, when an official “second” $\Delta GDP(I)$ estimate is available, the results using “second” growth rates are very similar to those reported in this section using the “third” growth rates. And, as discussed above, the results using the first annual revision growth rates are even more favorable to $\Delta GDP(I)$ than the results using “third” growth rates, showing a statistically significant tendency for $\Delta GDP(E)$ to revise towards $\Delta GDP(I)$ over the full sample.

For the “advance” estimates, when an official $\Delta GDP(I)$ is not available, the situation is quite different. It should be noted that the constructed “advance” $\Delta GDP(I)$ used in the Appendix employ only lags and other “advance” NIPA components to forecast profits; some companies have reported their quarterly profits numbers at the time of the “advance” release, and incorporating this information may produce a much-improved “advance” $\Delta GDP(I)$ estimate. That said, these rather limited “advance” $\Delta GDP(I)$ estimates perform poorly compared to the official “advance” $\Delta GDP(E)$ estimates, which better predict most of the business cycle variables used in this section. In addition, when predicting latest $\Delta GDP(I)$, about two-thirds weight should be placed on “advance” $\Delta GDP(E)$, and only about one-third weight should be placed on the constructed “advance” $\Delta GDP(I)$ estimates. This suggests that the initial estimates of corporate profits produced by the BEA are highly informative, and cannot be easily predicted based on lags or other available NIPA variables. For fourth quarter “second” estimates, when official profits numbers remain unavailable, this is presumably the case as well.

4 Information Content of the Latest Growth Rates

4.1 The Cyclicalities of the Latest Estimates

Table 1 showed that the correlation between $\Delta GDP(E)$ and $\Delta GDP(I)$ dropped sometime around the mid-1980s, and the divergences between the estimates also became highly cyclical around that time. Figure 5 shows this using year-over-year growth rates: GDP(I) rose faster than GDP(E) through most of the 1990s boom and the comparatively-short boom period from 2004 to 2006, while GDP(I) growth fell below GDP(E) growth in the 2001 recession and the latest cyclical downturn.⁷ Figure 6 plots the statistical discrepancy (GDP(E) minus GDP(I)) as a percent of GDP(E) versus the unemployment rate; work by Charles Fleischman first examined this relation, to my knowledge. Fleischman and John Roberts (2010) have studied the relation between GDP(E), GDP(I), the unemployment rate and other variables in the context of a state space model of the business cycle; their work points to the unemployment rate as an excellent measure of the state of the business cycle, as well as suggesting GDP(E) is measured with more error than GDP(I). Figure 6 shows the measurement errors in either GDP(I) or GDP(E) are clearly systematically related to the business cycle, and the statistical discrepancy is not noise, as is commonly assumed.

To understand the relation in Figure 6 better, consider a very simple model. It should be noted that Nalewaik (2008) shows why the type of model outlined below is an incomplete characterization of the growth rates of GDP(E) or GDP(I), and outlines models that fit the evidence better. However, the model outlined below is useful for the limited purpose of framing the subsequent discussion. Let true output be Y_t^* , and assume we can decompose this into trend τ_t and cycle ψ_t , so $Y_t^* = \tau_t + \psi_t$. The unemployment rate U_t is governed by

⁷We should keep in mind that these data are subject to further annual and benchmark revisions.

an Okun's law relation:

$$U_t - U_t^n = \gamma (Y_t^\star - \tau_t) = \gamma \psi_t, \quad \gamma < 0.$$

Now, assume GDP(I) and GDP(E) are systematically either too cyclical or not cyclical enough, so:

$$GDP(E)_t = \tau_t + \alpha_E \psi_t \quad \text{and:}$$

$$GDP(I)_t = \tau_t + \alpha_I \psi_t.$$

Then the statistical discrepancy $SD_t = GDP(E)_t - GDP(I)_t = (\alpha_E - \alpha_I) \psi_t$, and assuming the systematic mismeasurement is not identical for the two estimates, we should observe a relation between the discrepancy and the unemployment rate:

$$U_t - U_t^n = \frac{\gamma}{\alpha_E - \alpha_I} (SD_t) \quad \gamma < 0.$$

The strong positive relation shown in Figure 6 then implies $\alpha_E < \alpha_I$ - i.e. that the magnitude of the cycle is smaller in GDP(E) than in GDP(I). Table 3 shows regressions based on this relation. The first panel shows that the unemployment rate captures more than 60 percent of the variability of the discrepancy from 1984Q3 through 2006Q4, and though the statistical discrepancy is highly autocorrelated, the unemployment rate remains significant when an AR1 term is added. The second panel of table 3 shows specifications in first differences, to isolate the higher-frequency variation in the data. The first difference exhibits some negative autocorrelation, but the coefficient on the differenced unemployment rate remains positive, and the relation is highly significant when the differenced unemployment rate is lagged one quarter. These regression results confirm that the statistical discrepancy is not noise, even in differences.

Having established that $\alpha_E < \alpha_I$ within the context of this very stylized model, we can consider three possibilities:

1. Both GDP(I) and GDP(E) are more cyclical than true output, so $\alpha_I > \alpha_E > 1$. In this case, GDP(E) represents the cycle in true output better than GDP(I).
2. Both GDP(I) and GDP(E) are less cyclical than true output, so $\alpha_E < \alpha_I < 1$. In this case, GDP(I) represents the cycle in true output better than GDP(E).
3. GDP(E) is less cyclical and GDP(I) is more cyclical than true output, so $\alpha_E < 1 < \alpha_I$. In this case, GDP(I) represents the cycle in true output better than GDP(E) if $\alpha_I - 1 < 1 - \alpha_E$.

These possibilities frame the detailed discussion of the source data incorporated into the latest estimates in Appendix C. Plenty of evidence suggests GDP(E) misses part of the business cycle, implying possibility 1 is unlikely. Some of the construction components of GDP(E) are smoothed; in particular, the additions and alterations (adds and alts) component of residential structures is smoothed using a three-year moving average. This is problematic, because smoothed estimates inherently understate the magnitude of business cycle accelerations and decelerations. While adds and alts is a small component of GDP(E), it may have taken on outsized importance in the late 2000's downturn, and may have contributed to some of the fluctuations in the discrepancy around the 1990-1 recession. Probably more important, over most of this sample, the type of annual surveys used to compute the goods-producing sector of GDP(E) simply did not exist for most the (enormous) service-producing sector. As such, the BEA was forced to cobble together estimates based on trade-source, administrative, and regulatory data that may have missed part of the business cycle. For example, these data sources may miss fluctuations in the output of sole proprietors and some small businesses, highly cyclical parts of the economy. And the activities of many types of financial services companies or entities may have been missed

by the regulatory data used by the BEA to compute personal consumption expenditures (PCE) for financial services. The magnitude of the booms and busts in financial services, then, may not be fully reflected in the PCE component of GDP(E) or exports of services. However, many of these firms and entities likely did file tax forms, so their activities would have been represented in the tax data used to compute GDP(I). This could explain part of the increase in the statistical discrepancy in 1989, 2001, and in the latest episode.

Appendix C also discusses potential reasons why GDP(I) might be too cyclical. It is possible that some capital gains, which should be excluded from the BEA's definition of output, may have been misreported to the IRS as ordinary income and thus included in the tax data used to compute GDP(I). Capital gains are likely highly procyclical, so failure to exclude these gains could have made GDP(I) more cyclical than true output. Although the evidence on this is thin, possibility 3 might be slightly more likely than possibility 2. However, the evidence in favor of GDP(E) understating the cycle is stronger than the evidence in favor of GDP(I) overstating the cycle, so if possibility 3 holds, it is probably the case that $\alpha_I - 1 < 1 - \alpha_E$.

The last panel of table 3 shows that the statistical discrepancy is much less cyclical prior to the mid 1980s. Why might that be the case? While PCE for services has always held a relatively large share of GDP(E), averaging 30 percent from 1947 to 1984, its share shot up to an average of 43 percent from 1985 to 2009, and the share reached 48 percent in 2009. As the share of services PCE has increased, the measurement problems in GDP(E) may have become more severe, and more plainly visible. In addition, booms and busts in financial services may have accounted for a much larger share of the *variability* of the business cycle since the mid-1980s, with the junk bond boom and bust (as well as the savings and loan boom and bust) from the mid to late 1980s, the day-trading boom in the mid to late 1990s and subsequent stock market crash from 2000 to 2002, and the mortgage securitization boom and bust from 2002 to 2008. GDP(E) may have missed much of this

variation. But whatever the reason, since the really interesting divergences between the latest estimates occur in the post-1984Q3 period, the remainder of this section focuses on this sample.

4.2 Information in the Revisions about the Latest Estimates

Consider the following hypothetical example. We have two time series estimating the same unobserved variable of interest. The two time series happen to be identical, but we know the series are subject to considerable measurement error and may deviate a lot from the true variable of interest. Suppose new information becomes available that leads us to make large revisions to one of the estimates, bringing it closer to the truth, while the other estimate remains unrevised. Which estimate is now better? Obviously, the estimate that revised is better: we now know that it was way off initially and the revisions corrected some or all of that measurement error, and the estimate that did not revise remains way off. More generally, if the estimates start out identical, or pretty close, and the revisions improve the estimates, then the estimate that revises more, on average, will tend to be better than the estimate that revises less. This is the underlying logic of Fixler and Nalewaik (2007). Table 1 shows that the initial estimates of $\Delta GDP(I)$ and $\Delta GDP(E)$ do start out with a very high correlation, but $\Delta GDP(I)$ revises more. While the evidence in section 3 suggests that $\Delta GDP(I)$ starts out as the better estimate, if we make the relatively uncontroversial assumption that the revisions improve the estimates, then the larger revisions imply $\Delta GDP(I)$ expands on its lead. Fixler and Nalewaik (2007) use this revisions evidence to place bounds on the optimal weights to be placed on $\Delta GDP(I)$ and $\Delta GDP(E)$, and the bounds are favorable to $\Delta GDP(I)$.

The revisions increase the variance of $\Delta GDP(I)$ more than the variance of $\Delta GDP(E)$, implying that they add some news, or actual variation in true output growth, to $\Delta GDP(I)$ that is not added to $\Delta GDP(E)$ —see Mankiw, Runkle and Shapiro (1984), Mankiw and

Shapiro (1986), and Fixler and Nalewaik (2007). This variation in true output growth missed by latest $\Delta GDP(E)$ growth and captured by latest $\Delta GDP(I)$ is closely related to the business cycle. In particular, Nalewaik (2007b) shows that the revisions tend to reduce $\Delta GDP(I)$ more than $\Delta GDP(E)$ in low-growth states, so the extent of the weakness of true output growth in low-growth states appears to be part of the information missing from $\Delta GDP(E)$ but appearing in $\Delta GDP(I)$ through its more-informative revisions. Since this weakness in low-growth states appears in neither initial estimate and remains missing in latest $\Delta GDP(E)$, if latest $\Delta GDP(E)$ is correct, the revisions showing this relative weakness in latest $\Delta GDP(I)$ must damage the estimates. More broadly, any suggestion that latest $\Delta GDP(E)$ is better than latest $\Delta GDP(I)$ would seem to imply that the variability added to $\Delta GDP(I)$ through the revisions moves it further away from the true output growth. This seems hard to believe, and carried to its logical conclusion, the BEA should stop revising $\Delta GDP(I)$, and allocate its resources elsewhere. I do not think anyone at the BEA would seriously advocate taking that step. In contrast, the standard interpretation of the revisions is less problematic for the BEA: the revisions improve both $\Delta GDP(E)$ and $\Delta GDP(I)$, but the source data incorporated into $\Delta GDP(E)$ are just not as informative as the source data incorporated in $\Delta GDP(I)$. But in that case, latest $\Delta GDP(I)$ is likely the better estimate.

4.3 Relation to Other Business Cycle Variables

The logic behind these tests is similar to the logic behind the regression results in table 2, but table 4 switches the regression order and reports results from pairs of regressions, one of latest $\Delta GDP(I)$ and one of latest $\Delta GDP(E)$ on each cyclical indicator. The cyclical indicator is reported in the first column of the table, while the R^2 s in the second and third columns show that latest $\Delta GDP(I)$ is more highly correlated with every single one of these cyclical indicators. Appendix D repeats these results using annual instead of

quarterly data, and the results are quite similar.⁸ It is more highly correlated with lagged stock price changes, the lagged slope of the yield curve, and the lagged spread between high-yield corporate bonds and treasury bonds (using a somewhat shortened sample).⁹ It is more highly correlated with short and long differences of the unemployment rate, both contemporaneously and at leads and lags; the same holds true for the household survey measure of employment growth. Recall that there is no reason to suspect these measures to be spuriously correlated with $\Delta GDP(I)$; see footnote 6. It is more highly correlated with the manufacturing ISM, and using a shorter sample, the non-manufacturing ISM. It is also more highly correlated with dummies for NBER recessions; see also Nalewaik (2007a).¹⁰

As in table 2, latest $\Delta GDP(I)$ may be more highly correlated with all these variables because latest $\Delta GDP(E)$ is contaminated with more noise. But the interpretation of the revisions provided in the previous subsection suggests an alternative measurement error story, namely, that latest $\Delta GDP(E)$ misses variation in true output growth that appears

⁸Much of the source data incorporated at annual revisions is annual frequency, with no information on quarterly patterns, so the quarterly numbers are likely less reliable than the annuals. For example, the BEA is confident that employee gains from exercising nonqualified stock options net out of the annual GDP(I) estimates (since profits fall by the same amount as the increase in compensation), but they are concerned the quarterly pattern within years may be distorted.

⁹In his comments, J. Steven Landefeld suggests that the stock market may be more highly correlated with $\Delta GDP(I)$ because capital gains may be “leaking” into $\Delta GDP(I)$. If this were the case, the correlation between changes in the stock market and $\Delta GDP(I)$ should be contemporaneous, especially at the annual frequency, since a rising stock market translates immediately into a capital gain. Appendix D shows that the evidence does not support this: using the annual output growth measures, $\Delta GDP(E)$ is slightly more correlated with the contemporaneous change in the stock market, while $\Delta GDP(I)$ is more highly correlated with the stock market change from one year earlier. The evidence is more suggestive of either the stock market anticipating changes in true output, or changes in the stock market affecting true output with a lag, with true output better represented by $\Delta GDP(I)$. See also Nalewaik (2008).

¹⁰Note that these higher correlations, often substantially higher, are evidence against the crude model outlined in section 4.1. In that model, the $\Delta GDP(E)$ and $\Delta GDP(I)$ contain rescaled versions of the same business cycle fluctuations, in which case the R^2 must be equal across the two regressions. That is clearly not the case; $\Delta GDP(I)$ contains different business cycle fluctuations, fluctuations that also show up in these other business cycle variables. Nalewaik (2008) uses essentially this same argument to reject a crude rescaling model in favor of the LoSE model; see below. Nevertheless, both the LoSE model and the rescaling model say the same thing, broadly speaking: GDP(E) growth misses some of the business cycle fluctuations in true output growth, fluctuations that show up in GDP(I) growth as well as other variables.

in all these cyclical indicators and is also picked up by latest $\Delta GDP(I)$. In either case, latest $\Delta GDP(I)$ is the better estimate of true output growth. Is there any interpretation of these results where that is not the case? Latest $\Delta GDP(I)$ would have to contain measurement errors uncorrelated with true output growth, but correlated with all these other variables. That seems highly unlikely.

For the more econometrically-oriented reader, these regressions provide provide formal tests of what I consider the most likely hypothesis explaining these correlations (partially based on the revisions evidence in the previous subsection): that latest $\Delta GDP(E)$ is missing some of the variability of true output growth that is reflected in latest $\Delta GDP(I)$ and these other time series. Nalewaik (2008) derives such tests for Lack of Signal Error (LoSE), as the paper puts it. The maintained assumption is that the time series used for testing captures some of the variation missing from one estimate but included in the other. Regressions are run of each estimate on the testing variable. Nalewaik (2008) shows that the LoSE biases the regression coefficient on the testing variable towards zero, so the regression using the estimate that contains more LoSE yields a coefficient closer to zero. Note that it is measurement error of the LoSE form in the *dependent* variable that causes this attenuation bias, precisely opposite the conventional wisdom about classical measurement error (i.e. that it is measurement error in the *explanatory* variable that causes attenuation bias). Testing the equality of the coefficients on the testing variable across the two regressions, Nalewaik (2008) rejects using the asset price variables that are also employed in the first five specifications of table 4, using a slightly different sample. If the model assumptions hold, then $\Delta GDP(I)$ contains more signal about true output growth than $\Delta GDP(E)$, signal that is reflected in stock and bond prices. Table 4 shows that this missing signal also appears in the differenced unemployment rate, household survey employment growth, and the ISM measures. The coefficients in table 4 are all larger, in absolute value, when $\Delta GDP(I)$ is the dependent variable; a relatively large amount of noise in $\Delta GDP(E)$ cannot explain

these results, but a relatively large amount of LoSE can.

Less formal comparisons of GDP(E) and GDP(I) with other sources of information about the business cycle are also informative. In particular, we can compare the peaks and troughs in GDP(E) and GDP(I) with the NBER peak and trough dates. Grimm (2005) does this; Figures 7a, 7b, and 7c show the results graphically for the three recessions prior to the most recent one. The one case where GDP(I) seems to differ from the NBER is the 1990-1 recession: GDP(I) starts declining during the NBER peak quarter while GDP(E) is flat, but since the monthly peak was July 1990, the 1990Q3 GDP(I) decline seems consistent with the NBER dating. In the 1981-2 recession, the NBER called the trough in 1982Q4, the same quarter as GDP(I), while GDP(E) calls the trough three quarters earlier. In the 2001 recession, it is difficult to discern any real cyclical downturn in GDP(E), while the NBER peak and trough dates line up perfectly with GDP(I). These peak and trough dates summarize the information in several other reliable indicators, and the fact that they line up better with GDP(I) is again suggestive that GDP(I) is the better estimate.

5 The Estimates over the 2007-2009 Cyclical Downturn

The recent downturn looks considerably worse when output is measured using GDP(I) instead of GDP(E). First, the effect on output appears sooner, with GDP(I) showing a sharp deceleration even before the NBER peak in late 2007. This deceleration was somewhat evident in the real time estimates of GDP(I), but more importantly, the actual recession itself was much more evident in the real time estimates of GDP(I) than in the real time estimates of GDP(E). Second, the steepness of the plunge in output in late 2008 and early 2009 appears worse. And third, with the BEA's February 2010 data release, the decline in output now appears more prolonged, extending into the summer of 2009.

Figure 8 shows levels of GDP(E) and GDP(I) as they were measured at different dates, indexing to 2006Q1 because the levels of the two series are different and have changed with revisions. The light and dark dotted lines show the GDP(E) and GDP(I) estimates towards the start of the recession, in early 2008. The estimates are trending up at a similar pace through 2006 and the first half of 2007, but the estimates then diverge considerably. GDP(I) shows an economy in a much more vulnerable state in late 2007, with output essentially flat over the second half of 2007. Meanwhile, GDP(E) showed little of this vulnerability in the second half; although growth was weak in 2007Q4, that weakness came on the heels of estimated annualized growth of almost 5 percent in 2007Q3.¹¹

The light and dark dashed lines of Figure 8 show the GDP(E) and GDP(I) estimates at the end of December 2008, after the NBER had called December 2007 a business cycle peak. All four of the monthly indicators the NBER uses to date business cycles had peaked in late 2007 and early 2008, and GDP(I) was trending down slightly through the first three quarters of 2008 as well. GDP(E) was the only anomaly, showing continued growth at an annual rate of almost 2 percent in the first half of 2008.

The light and dark solid lines of Figure 8 show the latest GDP(E) and GDP(I) estimates, and we see that the initial $\Delta GDP(E)$ estimates for 2008 have since revised down considerably towards the initial $\Delta GDP(I)$ estimates, a notable continuation of the recent pattern in revisions discussed in section 3. Revisions have also reduced GDP(I), but the revisions to growth came mainly in the first half of 2007. The latest estimates show that

¹¹At the end of March 2008, the bivariate Markov switching model using $\Delta GDP(I)$ in Nalewaik (2007a) estimated a probability of around 90% that the economy had downshifted to a low growth state by 2007Q4, and this probability remained well above 50% throughout 2008. At the same time, a Markov-switching model using $\Delta GDP(E)$ alone estimated a probability of less than 20% that the economy had downshifted to a low-growth state, a probability that remained low through most of 2008 (for example 27% at the end of September), only cracking 50% after the BEA's "advance" 2008Q3 estimates released at the end of October. Models based on monthly indicators did not do better: an implementation of the Diebold-Rudebusch (1996) monthly indicators model, based on Kim and Nelson (2000), did not jump above 50 percent until early November 2008, with the BEA's release of its initial personal income numbers for September 2008. The behavior of these models shows that real time assessments of the state of the business cycle can be meaningfully improved by looking at GDP(I).

GDP(I) was essentially flat over the four quarters of 2007, declining in 2007Q1 and 2007Q3. These latest estimates suggest the recent cyclical downturn caused a measurable deceleration in aggregate output much earlier than is commonly believed. Meanwhile, GDP(E) currently shows no such early deceleration, growing 2.5 percent over the four quarters of 2007, about the same as in 2006. These differences over 2007 produce the bulk of the enormous swing in the statistical discrepancy we saw in Figure 6, from around minus 1.9 percent of GDP(E) in late 2006 to plus 1.8 percent of GDP(E) in 2009Q3.

The current estimates show $\Delta GDP(E)$ actually slightly weaker than $\Delta GDP(I)$ in the first three quarters of 2008, but the current estimates of $\Delta GDP(I)$ then show a steeper downturn over the worst part of the recession. The current annualized $\Delta GDP(I)$ estimates for 2008Q4 and 2009Q1 are -7.3 and -7.7 percent, worse than the $\Delta GDP(E)$ estimates of -5.4 and -6.4 percent.

Finally, the latest $\Delta GDP(I)$ estimates for 2009Q3, released in late February 2010 and incorporating numbers from the Quarterly Census of Employment and Wages (see Appendix C), have called into question the timing of the trough of the recession. Before these numbers were released, a conventional wisdom was emerging that the recession likely ended late in the second quarter of 2009, perhaps in June, with the economy resuming growth in 2009Q3. Figure 8 shows a modest rebound in GDP(E) in 2009Q3, but no evidence of a rebound in GDP(I). Personal income less transfer payments and employment—two of the four indicators most emphasized by the NBER business cycle dating committee—have continued to decline in 2009Q3.

What are we to make of these important differences between GDP(E) and GDP(I) over this cycle? It should be noted that all these estimates remain subject to considerable future revision, but the source data is most concrete for 2007, which happens to be the period of the greatest widening of the statistical discrepancy. Currently, the corporate profits and proprietors' income components of GDP(I) incorporate IRS tax returns data through 2007,

and declines in these two income categories account for the bulk of the deceleration in GDP(I) that year. Proprietors' income increased about \$63 billion (nominal) in 2006 and fell \$37 billion in 2007,¹² a deceleration of about \$100 billion. The biggest declines in 2007 were in real estate, construction, finance and insurance, as well as (less explicably) mining; see BEA table 6.12D. As noted in Appendix C, it is possible that some of the decline in proprietors' income may have represented a decline in capital gains from house flipping, which should not be included in the relevant concept of output. Real estate proprietors' income fell \$24 billion in 2007, but it also fell \$14 billion in 2006, suggesting this type of mismeasurement cannot explain much of the widening of the statistical discrepancy in 2007. Construction proprietors' income decelerated from a \$6 billion increase in 2006 to a \$14 billion decline in 2007, with the current estimates showing a \$46 billion decline in 2008. Part of this decline in proprietors' income should probably have shown up in lower spending on residential improvements, but as discussed earlier and in Appendix C, the BEA's averaging of their raw source data will tend to miss such a large deceleration. Currently, the raw estimates of improvements spending from Census show 4 and 14 percent declines in 2007 and 2008, respectively, steeper than the current BEA estimates of 1 and 4 percent declines. If the Census numbers are correct, GDP(E) should be \$5 billion lower in 2007 and about \$22 billion lower in 2008, so this also explains only a small portion of the widening of the statistical discrepancy.¹³

¹²The 2007 decline in the raw IRS tax numbers was larger, about \$66 billion (see BEA table 7.14), but the BEA cut this down with various adjustments, including the inventory valuation and capital consumption adjustments.

¹³Some other data sources suggest much larger declines in spending on residential improvements. For example, Greenspan and Kennedy (2005, 2007) use Flow of Funds data and Mian and Sufi (2009) use data from credit rating agencies to show that households extracted a very large amount of home equity in the mid-2000s, before banks cut credit lines in 2007 and 2008 and equity extraction dropped dramatically. Using survey evidence that households spend about a third of extracted home equity on home improvements—see Brady, Canner, and Maki (2000), Canner, Dynan and Passmore (2002), and Greenspan and Kennedy (2007) and the references therein—updated Greenspan-Kennedy estimates give declines in spending on home improvements of \$65 billion in 2007 and \$80 billion in 2008. Of course, this does not necessarily imply causality from equity extraction to spending, because households may have found other financing

Corporate profits' increased about \$152 billion in 2006, and fell \$67 billion in 2007 (a deceleration of about \$220 billion), and the current estimates for 2008 show a decline of \$181 billion. The biggest decline in profits in 2007 was in the finance and insurance industry; the \$54 billion decline in 2007 followed an increase of \$4 billion in 2006 and a massive increase of about \$180 billion from 2000 to 2005. Looking more broadly, the sum of corporate profits, proprietors' income, and wage and salaries for the finance and insurance industry fell close to 4 percent in 2007, while PCE for financial services increased more than 12 percent. While these categories are not strictly comparable, this is difficult to reconcile without severe measurement error in either the income measures or PCE.¹⁴ The BEA started worrying back in late 2007 about their ability to strip out capital losses (bad debt expenses and asset write-downs) from their initial estimates of financial companies' profits, but the availability of the tax data for 2007 likely made these subtractions much easier. With the tax data ameliorating this issue, the problems appear more concentrated in the measurement of financial services PCE and services more generally on the expenditure side, as discussed in the previous section and in Appendix C. Given the advent of the financial crisis and the disappearance of many securitization markets in the second half of 2007, a 12 percent growth rate for financial services PCE seems implausibly high. To get a sense of the magnitudes involved, a decline in financial services PCE of 4 percent would have lowered GDP(E) in 2007 by \$76 billion from its current level, and by more if PCE missed the boom in financial services output over prior years. More recently, profits in the finance and insurance industry fell an additional \$91 billion in 2008 (with proprietors' income and wage and salary income also falling), while financial services PCE increased once again. Since the tax data have not yet been incorporated for 2008, some risk remains that the

options absent the availability of home equity lines of credit.

¹⁴The output of financial services could also have shown up in exports, or as an intermediate input into the production of other industries.

income declines were too steep, but again it seems implausible that financial services PCE continued its uninterrupted growth.

Overall, this evidence suggests that although there may be problems on both sides of the accounts, the problems are likely more severe on the expenditure side. Given that, the latest downturn was likely substantially worse than the current GDP(E) estimates show. Output likely decelerated sooner, fell at a faster pace at the height of the downturn, and recovered less quickly than is reflected in GDP(E), and in conventional wisdom.

6 Concluding Thoughts

Considerable evidence suggests that the growth rates of GDP(I) better represent the business cycle fluctuations in true output growth than do the growth rates of GDP(E). For the initial growth rates, the revisions evidence over the past 15 years, the correlations with other business cycle indicators, and the recent behavior of the estimates around cyclical turning points all point to this conclusion. For the latest estimates that have passed through their cycle of revisions, careful consideration of the nature of the source data, statistical analysis of the information added by the revisions, and statistical tests as well as informal comparisons with other business cycle indicators, again all suggest GDP(I) growth is better than GDP(E) growth.

These results strongly suggest that economists and statisticians interested in business cycle fluctuations in U.S. output should pay attention to the income-side estimates, and consider using some sort of weighted average of the income- and expenditure-side estimates in their analyses. The evidence in this paper clearly suggests that the weights should be skewed towards GDP(I), and Fixler and Nalewaik (2007) are able to place fairly tight bounds on the optimal weights for the latest estimates, bounds favorable to GDP(I). But a 50-50 average would be a marked improvement over an average placing all its weight on

GDP(E). Such a 50-50 average would be following in the footsteps of the Council of Economic Advisers, who, after concluding GDP(I) might be better than GDP(E) in their 1997 Economic Report of the President, have subsequently given some weight to the income-side estimates in their productivity analyses—see the 2008 Economic Report of the President, p. 39, and the 2009 Economic Report of the President, pp. 47-8.

The results here also have implications for the BEA. When a quarterly estimate of GDP(I) growth is available, the evidence here shows it is likely a better estimate of output growth than the corresponding GDP(E) estimate. However, the first GDP(E) estimate for any given quarter, the “advance” estimate, is typically released about a month before the first GDP(I) estimate, and GDP(I) is delayed an additional month when the BEA is producing estimates for fourth quarters. These delays occur because the BEA has incomplete information on corporate profits, and is not comfortable releasing earlier estimates of profits. In general, the profits information released by the BEA appears tremendously useful, and the BEA does have some information on profits at these earlier release dates. An “advance” estimate of GDP(I) based on the available profits information might be quite helpful for real time assessment of the speed of economic growth. Earlier release of the fourth-quarter estimates, so a GDP(I) estimate is released at least as early as the BEA’s “second” release, might be similarly helpful; the BEA has still not released an estimate of GDP(I) growth for the fourth quarter of 2009 at the time of this paper, March 19th, 2010. These decisions will depend on how much information on profits is really available at these earlier dates, and a thorough assessment of this issue seems to be in order.

The BEA, the Census Bureau, and the BLS doubtless will continue making improvements in their estimates where feasible, and the good news is that there have been substantial improvements over the past few years. The data on services have taken leaps and bounds forward, with the advent of Census’s Quarterly Service Survey in 2003, the recent expansions in the coverage of this survey, and the expansions in the coverage of the Service

Annual Survey. Further improvements are in train: in December 2010, the estimates from the Service Annual Survey will roughly double in coverage, expanding to mimic the sector coverage of the Economic Census.¹⁵ These data should improve the estimates of PCE and GDP(E).

However, despite these improvements, problems with the output growth estimates will inevitably remain, and lack of coverage of services is only one of several important limitations of GDP(E). All the results in this paper suggest that the current reporting practice of the BEA, which puts nearly exclusive emphasis on GDP(E) over GDP(I), is suboptimal statistically. The BEA creates tremendous value by producing an income-based estimate of output growth, but the current BEA reporting practice downplays it so much that many analysts may not even be aware of its existence. The BEA's typical press release hardly ever discusses GDP(I), and it is reported only towards the back of the release as a nominal level, so analysts would have to deflate and compute annualized quarterly growth rates themselves to arrive at the number comparable to headline real GDP(E) growth.

If the BEA found the results here persuasive, they may want to consider taking several incremental steps towards increasing the prominence of GDP(I). Most obviously, the BEA could report real annualized growth rates of GDP(I) in its press releases, preferably in table 1 of the release so they can be compared easily to the annualized growth rates of real GDP(E). Second, it could give those annualized growth rates more prominence in the text of the press releases, discussing them at a level of detail similar to its current discussion of GDP(E). The BEA's discussion of the corporate profits estimates could be rolled into a more general discussion of GDP(I). Third, the BEA could bring more balance to their statements about the reliability of GDP(E) and GDP(I). Landefeld, Seskin and Fraumeni (2008) take a small step in this direction by stating: "... these studies remind users that it is useful to look at growth in both GDP and gross domestic income in assessing the current

¹⁵<http://www.census.gov/services/index.html>

state of the economy.”¹⁶

Featuring two measures of output growth in the same press release would raise communication challenges, and the BEA might fear that having two featured measures could be too confusing for casual analysts.¹⁷ If the BEA is concerned about the communication issue, they may look to the example of other countries such as Great Britain and Australia, who report an average estimate from the different sides of the accounts as their featured output growth measure. The BEA has considered taking this step in the past—see, for example, Moulton (2000)—and the BEA could report such an average of GDP(E) and GDP(I) as GDP(A).¹⁸ The BEA could employ optimal weights guided by statistical analysis, as in Fixler and Nalewaik (2007), but the results here suggest that featuring even a straight 50-50 average would be a marked improvement over the current practice of featuring GDP(E) alone.

References

- [1] Abel, A. and Bernanke, B. (2001) *Macroeconomics*, Fourth Edition, Boston: Addison Wesley.

¹⁶Landefeld, Seskin and Fraumeni (2008), p. 211.

¹⁷The BLS does reports two estimates of the monthly change in employment in its employment report, one from both the establishment survey and one from the household survey, but there are clear statistical reasons for favoring the establishment survey number at the monthly frequency. For the case of GDP(E) and GDP(I), making the case in favor of one measure over the other is more complicated.

¹⁸Of course, the components of GDP(E) will not sum to the topline GDP(A), nor will the components of GDP(I) sum to the topline, which may be confusing for some analysts. But if the evidence in this paper is convincing, it is the current state of affairs that the components of GDP(E) do not sum to true output or even the best estimate of true output; in fact, the sum of the components of GDP(E) miss important, systematic variation in true output. The new reporting will simply make these facts explicit. Over the long-run, allocating parts of the discrepancy to different components of GDP(E) and GDP(I) may be the right thing to do, but it is an extremely complicated task, and much research will need to be done before implementing anything. But if BEA attempts to go down this path at some point, they should do this in a transparent and easily replicable fashion. The BEA is to be commended for their transparency in reporting the statistical discrepancy, and should do nothing to compromise this transparency.

- [2] Beaulieu, J. and Bartelsman, E. (2004) “Integrating Expenditure and Income Data: What to do with the Statistical Discrepancy?” Board of Governors of the Federal Reserve System, FEDS working paper 2004-39.
- [3] Brady, P., Canner, G., and Maki, D. (2000), “The Effects of Recent Mortgage Refinancing.” *Federal Reserve Bulletin* July 2000: 441-450.
- [4] Bureau of Economic Analysis, U.S. Department of Commerce (2002) “Corporate Profits,” Methodology Paper, September 2002.
- [5] Bureau of Economic Analysis, U.S. Department of Commerce (2008) “Updated Summary of NIPA Methodologies,” *Survey of Current Business*, November 2008, pp. 8-25.
- [6] Bureau of Economic Analysis, U.S. Department of Commerce (2008) “Concepts and Methods of the U.S. National Income and Product Accounts (Introductory Chapters 1-4),” July 2008.
- [7] Council of Economic Advisers (1997), *Economic Report of the President*, Washington, D.C.: United States Government Printing Office.
- [8] Council of Economic Advisers (2009), *Economic Report of the President*, Washington, D.C.: United States Government Printing Office.
- [9] Canner, G., Dynan, K., and Passmore, W. (2002), “Mortgage Refinancing in 2001 and Early 2002.” *Federal Reserve Bulletin* December 2002: 469-481.
- [10] Diebold, Francis X. and Rudebusch, Glenn D. “Measuring Business Cycles: A Modern Perspective.” *Review of Economics and Statistics*, 1996 (101), pp. 67-77.
- [11] Dynan, K. and Elmendorf, D. (2001), “Do Provisional Estimates of Output Miss Economic Turning Points?” Board of Governors of the Federal Reserve System, FEDS working paper 2001-52.
- [12] Fixler, D. and Grimm, B. (2006) “GDP Estimates: Rationality tests and turning point performance,” *Journal of Productivity Analysis*, 25, pp. 213-229.
- [13] Fixler, D. and Nalewaik, J. (2007) “News, Noise, and Estimates of the True Unobserved State of the Economy,” Board of Governors of the Federal Reserve System, FEDS working paper 2007-34.
- [14] Fleischman, C. and Roberts, J. (2010) “A Multivariate Estimate of Trends and Cycles,” Federal Reserve manuscript.
- [15] Greenspan, A. (2004) “Risk and Uncertainty in Monetary Policy” *American Economic Review*, 94, pp. 33-40.
- [16] Greenspan, A. and Kennedy, J. (2005) “Estimates of Home Mortgage Originations, Repayments, and Debt on One-to-Four-Family Residences,” Board of Governors of the Federal Reserve System, FEDS working paper 2005-41.

- [17] Greenspan, A. and Kennedy, J. (2007) "Sources and Uses of Equity Extracted from Homes" Board of Governors of the Federal Reserve System, FEDS working paper 2007-20.
- [18] Grimm, B. (2005) "Alternative Measures of U.S. Economic Activity in Business Cycles and Dating," BEA working paper 2005-05.
- [19] Grimm, B. and Weadock, T. (2006), "Gross Domestic Product, Revisions and Source Data." *Survey of Current Business* February 2006: 11-15.
- [20] Holdren, A. and Grimm, B. (2008), "Gross Domestic Income, Revisions and Source Data." *Survey of Current Business* December 2008: 14-20.
- [21] Hamilton, James D. 1990. Analysis of Time Series Subject to Changes in Regime. *Journal of Econometrics* 45: 39-70.
- [22] Koenig, Evan; Dolmas, Sheila; and Piger, Jeremy. "The Use and Abuse of Real-Time Data in Economic Forecasting." *Review of Economics and Statistics* 85 (August 2003), pp. 618-628.
- [23] Krakower, H. (2007) "Improved Measures of the Misreporting of Income," presentation at National Economic Accounts Data Users' Conference, April 2007.
- [24] Landefeld, E., Seskin, E., and Fraumeni, B. "Taking the Pulse of the Economy: Measuring GDP." *Journal of Economic Perspectives* 22 (Spring 2008), pp. 193-216.
- [25] Mankiw, N., Runkle, D., and Shapiro, M. "Are Preliminary Announcements of the Money Stock Rational Forecasts?" *Journal of Monetary Economics*, 1984 (14), pp. 15-27.
- [26] Mankiw, N. and Shapiro, M. "News or Noise: An Analysis of GNP Revisions" *Survey of Current Business*, May 1986, pp. 20-25.
- [27] McCully, C., and Payson, S. (2009), "Preview of the 2009 Comprehensive Revision of the NIPAs." *Survey of Current Business* May 2009: 6-16.
- [28] Mian, A. and Sufi, A. (2009) "House Prices, Home Equity-Based Borrowing, and the U.S. Household Leverage Crisis," manuscript, University of Chicago Booth School of Business, July 2009.
- [29] Moulton, B. (2000) "Getting the 21st-Century GDP Right: What's Underway?" *American Economic Review*, 90, pp. 253-258.
- [30] Moylan, C. (2008) "Employee Stock Options and the National Economic Accounts," *Survey of Current Business*, February 2008, pp. 7-13.
- [31] Nalewaik, J., (2007a), "Estimating Probabilities of Recession in Real Time Using GDP and GDI," Board of Governors of the Federal Reserve System, FEDS working paper 2007-07.

- [32] Nalewaik, J., (2007b), “Incorporating Vintage Differences and Forecasts into Markov Switching Models.” Board of Governors of the Federal Reserve System, FEDS working paper 2007-23.
- [33] Nalewaik, J., (2008), “Lack of Signal Error (LoSE) and Implications for OLS Regression: Measurement Error for Macro Data.” Board of Governors of the Federal Reserve System, FEDS working paper 2008-15.
- [34] Perry, G. (2005), “Gauging Employment: Is the Professional Wisdom Wrong?” *Brookings Papers on Economic Activity* 2005(2): 285-321.
- [35] Seskin, E., and Smith, S. (2009), “Improved Estimates of the National Income and Product Accounts, Results of the 2009 Comprehensive Revision.” *Survey of Current Business* September 2009: 15-35.

**Table 1: Variances and Correlations,
Initial (3rd) and Latest Estimates of Δ GDP(E) and Δ GDP(I)**

Correlation Matrix, 1978Q3 to 2009Q3:

	Initial (3rd) Estimates		Latest Estimates	
	Δ GDP(E)	Δ GDP(I)	Δ GDP(E)	Δ GDP(I)
Init. $\Delta GDP(E)$	1.00	0.95	0.85	0.77
Init. $\Delta GDP(I)$	0.95	1.00	0.81	0.82
Latest $\Delta GDP(E)$	0.85	0.81	1.00	0.79
Latest $\Delta GDP(I)$	0.77	0.82	0.79	1.00

Correlation Matrix, 1984Q3 to 2006Q4:

	Initial (3rd) Estimates		Latest Estimates	
	Δ GDP(E)	Δ GDP(I)	Δ GDP(E)	Δ GDP(I)
Init. $\Delta GDP(E)$	1.00	0.90	0.68	0.63
Init. $\Delta GDP(I)$	0.90	1.00	0.61	0.66
Latest $\Delta GDP(E)$	0.68	0.61	1.00	0.60
Latest $\Delta GDP(I)$	0.63	0.66	0.60	1.00

Variances:

1978Q3 to 2009Q3		1984Q3 to 2006Q4	
Δ GDP(E)	Δ GDP(I)	Δ GDP(E)	Δ GDP(I)
Initial (3rd) Estimates			
8.53	8.90	3.88	3.89
Latest Estimates			
9.44	10.29	4.23	4.96
Revision: Latest-Initial (3rd)			
2.78	3.60	2.57	3.05

Table 2: Predictive Content of Initial Growth Rates, 1978Q3-2009Q3

Dependent Variable	Explanatory Variables						Constant	Adj. R^2
	$\Delta GDP(E)$, initial (3rd)			$\Delta GDP(I)$, initial (3rd)				
	t	$t-1$	$t-2$	t	$t-1$	$t-2$		
$(GDP(E)_t/GDP(E)_{t-1})^4$, latest	0.77			0.12			0.40	0.71
	(0.13)			(0.12)			(0.24)	
	0.76	-0.17		0.07	0.27		0.28	0.72
	(0.12)	(0.15)		(0.11)	(0.15)		(0.21)	
$(GDP(E)_t/GDP(E)_{t-1})^4$, initial (3rd)		-0.46			0.93		1.42	0.25
		(0.28)			(0.29)		(0.32)	
		-0.37			0.80		1.38	0.22
		(0.26)			(0.25)		(0.28)	
$(GDP(I)_t/GDP(I)_{t-1})^4$, latest			-0.33			0.54	1.97	0.05
			(0.22)			(0.26)	(0.33)	
	0.02			0.86			0.34	0.66
	(0.18)			(0.18)			(0.25)	
$(GDP(I)_t/GDP(I)_{t-1})^4$, initial (3rd)	0.03	-0.02		0.81	0.10		0.24	0.66
	(0.19)	(0.18)		(0.20)	(0.17)		(0.29)	
		-0.17			0.66		1.34	0.22
		(0.28)			(0.28)		(0.45)	
$(GDP(I)_t/GDP(I)_{t-1})^4$, initial (3rd)		-0.17		0.67			1.31	0.26
		(0.21)		(0.21)			(0.32)	
			-0.43			0.69	1.91	0.09
			(0.21)			(0.24)	(0.40)	
$(UR_t - UR_{t-1}) * 4$	0.03			-0.36			1.01	0.57
	(0.08)			(0.08)			(0.18)	
		0.02			-0.30		0.87	0.40
		(0.07)			(0.08)		(0.22)	
$(E_t^{household}/E_{t-1}^{household})^4$			0.05			-0.26	0.69	0.22
			(0.10)			(0.12)	(0.24)	
	0.08			0.35			0.06	0.47
	(0.12)			(0.11)			(0.27)	
$ISM_t^{manuf.}$		-0.03			0.41		0.17	0.38
		(0.10)			(0.13)		(0.33)	
			-0.19			0.43	0.55	0.16
			(0.13)			(0.14)	(0.39)	
$\log(SP500_t - SP500_{t-4})/4$	0.15			1.33			47.23	0.53
	(0.41)			(0.41)			(0.68)	
		0.41			0.93		47.59	0.43
		(0.38)			(0.37)		(0.66)	
$r_{t-8}^{Treas(10yr)} - r_{t-8}^{Treas(2yr)}$			-0.03			0.83	49.00	0.15
			(0.46)			(0.44)	(0.78)	
	-0.09			0.73			0.08	0.19
	(0.30)			(0.33)			(1.08)	
$\Delta \widehat{GDP}(E)_{t,t}$, SPF forecast current quarter	-0.14			0.21			0.51	0.08
	(0.07)			(0.08)			(0.15)	
	0.06			0.43			0.74	0.51
	(0.11)			(0.12)			(0.28)	
$\Delta \widehat{GDP}(E)_{t,t-1}$, SPF forecast 1 quarter ahead	0.03			0.22			1.68	0.19
	(0.07)			(0.09)			(0.30)	
	0.08			0.08			2.20	0.11
	(0.07)			(0.08)			(0.33)	
$(GDP(E)_t/GDP(E)_{t-1})^4$, latest, 1994Q1-2006Q4								
	0.58			0.42			-0.29	
	(0.16)			(0.16)			(0.20)	
	0.56			0.28			0.31	0.47
$(GDP(I)_t/GDP(I)_{t-1})^4$, latest, 1994Q1-2006Q4	(0.16)			(0.16)			(0.35)	
	0.15			0.85			-0.16	
	(0.24)			(0.24)			(0.26)	
	0.13			0.74			0.31	0.41
	(0.23)			(0.22)			(0.59)	

Table 3: Regressions Explaining Statistical Discrepancy as Percent of GDP(E)

Levels Specifications: $SD_t = \alpha + \beta SD_{t-1} + \theta U_t + \varepsilon_t$

1984Q3-2006Q4			
β	θ	α	Adj. R^2
0.93 (0.05)		-0.01 (0.05)	0.83
	0.88 (0.11)	-4.69 (0.64)	0.63
0.75 (0.12)	0.25 (0.11)	-1.36 (0.62)	0.84

Differences Specifications: $\Delta SD_t = \alpha + \beta \Delta SD_{t-1} + \theta_0 \Delta U_t + \theta_1 \Delta U_{t-1} + \varepsilon_t$

1984Q3-2006Q4				
β	θ_0	θ_1	α	Adj. R^2
-0.33 (0.17)			-0.04 (0.05)	0.10
-0.35 (0.18)	0.28 (0.23)		-0.03 (0.04)	0.10
-0.36 (0.16)		0.84 (0.25)	-0.01 (0.04)	0.22

Levels Specifications: $SD_t = \alpha + \beta SD_{t-1} + \theta U_t + \varepsilon_t$

1948Q1-1984Q2			
β	θ	α	Adj. R^2
0.74 (0.06)		0.15 (0.05)	0.55
	0.10 (0.05)	0.07 (0.28)	0.06
0.72 (0.06)	0.03 (0.02)	-0.02 (0.12)	0.55

Note: Standard errors are Newey-West with 8 lags.

Table 4: Regressions of $\Delta GDP(I)$ and $\Delta GDP(E)$ on Various Alternative Business Cycle Indicators, 1984Q3-2006Q4

Explanatory Variable	Adjusted R^2		β		p-val., equal β s
	$GDP(I)_t$	$GDP(E)_t$	$GDP(I)_t$	$GDP(E)_t$	
$\log(SP500_t/SP500_{t-7})/7$	0.14	0.08	0.29 (0.06)	0.20 (0.08)	0.03
$r_{t-1}^{HYcorporate} - r_{t-1}^{Treas.(7yr)}$	0.28	0.19	-0.67 (0.10)	-0.51 (0.14)	0.06
$r_{t-2}^{HYcorporate} - r_{t-2}^{Treas.(7yr)}$	0.20	0.11	-0.57 (0.13)	-0.41 (0.13)	0.02
$r_{t-3}^{HYcorporate} - r_{t-3}^{Treas.(7yr)}$	0.18	0.06	-0.54 (0.16)	-0.30 (0.15)	0.00
$r_{t-8}^{Treas.(10yr)} - r_{t-8}^{Treas.(2yr)}$	0.05	0.01	0.70 (0.37)	0.36 (0.38)	0.02
$(UR_t - UR_{t-1}) * 4$	0.26	0.24	-1.47 (0.27)	-1.32 (0.30)	0.40
$UR_t - UR_{t-4}$	0.25	0.10	-1.74 (0.29)	-1.04 (0.32)	0.00
$UR_{t+2} - UR_{t-2}$	0.35	0.21	-2.19 (0.28)	-1.59 (0.34)	0.02
$UR_{t+4} - UR_t$	0.24	0.18	-1.81 (0.31)	-1.46 (0.31)	0.08
$(E_t^{household}/E_{t-1}^{household})^4$	0.30	0.20	1.00 (0.17)	0.76 (0.21)	0.11
$E_t^{household}/E_{t-4}^{household}$	0.20	0.12	1.01 (0.19)	0.74 (0.21)	0.00
$E_{t+2}^{household}/E_{t-2}^{household}$	0.34	0.24	1.36 (0.19)	1.06 (0.22)	0.04
$E_{t+4}^{household}/E_t^{household}$	0.23	0.18	1.14 (0.23)	0.93 (0.20)	0.13
$ISM_t^{manuf.}$	0.33	0.19	0.29 (0.05)	0.20 (0.06)	0.01
$ISM_t^{non-manuf.}$	0.29	0.18	0.33 (0.08)	0.24 (0.05)	0.20
<i>Recession Dummies</i>	0.29	0.24	-5.05 (0.43)	-4.28 (0.77)	0.46

Note: Specifications using the high yield bond spread ($r^{HYcorporate} - r^{Treas.(10yr)}$) use a 1988Q3 to 2006Q4 sample. Specifications using the non-manufacturing ISM ($ISM^{non-manuf.}$) use a 1997Q3 to 2006Q4 sample. Standard errors are Newey-West with 8 lags.