News, Noise and Estimates of the "True" Unobserved State of the Economy

Dennis J. Fixler and Jeremy J. Nalewaik

WP2004-08
January 31, 2006

News, Noise, and Estimates of the "True" Unobserved State of the Economy

Dennis J. Fixler and Jeremy J. Nalewaik*

Draft, Comments Welcome

January 31, 2006

Abstract

Which provides a better estimate of the "true" state of the U.S. economy, gross domestic product (GDP) or gross domestic income (GDI)? Past work has assumed the difference between each estimate and the "true" state of the economy is pure *noise*, taking greater variability to imply lower reliability. However each difference may be pure *news* instead; then greater variability implies higher information content and greater reliability. We analyze various vintages of estimates, developing models for combining GDP and GDI under the differing assumptions, and use revisions to the estimates to show the news assumption is probably more accurate.

JEL classification: C1, C82.

Keywords: GDP, statistical discepancy, news and noise, signal-to-noise ratios, optimal combination of estimates, business cycles

*Fixler: U.S. Bureau of Economic Analysis, 1441 L Street NW, Washington, DC 20230 (e-mail: dennis.fixler@bea.gov); Nalewaik: U.S. Bureau of Economic Analysis, 1441 L Street NW, Washington, DC 20230 (e-mail: jeremy.nalewaik@bea.gov). We thank Erick Sager for research assistance, and Ben Bridgman, Ufuk Demiroglu, Bob Dennis, Bruce Grimm, Steve Landefeld, Keith Phillips, Matt Pritsker, Marshall Reinsdorf, Mark Watson, and seminar participants at the Bureau of Economic Analysis, the Washington Statistical Society, the Board of Governors of the Federal Reserve System, and the 2005 NBER SI Macroeconomics and Productivity workshop for comments. The views expressed in this paper are soley those of the authors and are not necessarily those of the U.S. Bureau of Economic Analysis or the U.S. Department of Commerce.

For analysts of economic fluctuations, estimating the true state of the economy from imperfectly measured official statistics is an ever-present problem. As most economists agree that no one statistic is a perfect gauge of the state of the economy, many have proposed using instead some type of weighted average of multiple imperfectly measured statistics. Examples include the composite index of coincident indicators, and averages of different measures of aggregate economic activity such as GDP and GDI. While the precise meaning of the state of the economy can vary from case to case, in this paper we fix ideas, taking it to mean the growth rate of the size of the economy as traditionally defined in the U.S. National Income and Product Accounts (NIPAs).[1]

The main point of our paper is as follows. Prior attempts to produce such a weighted average of imperfectly measured statistics have made a strong implicit assumption that drives their weighting: that the difference between the true state of the economy and each measured statistic is pure *noise*, or completely uncorrelated with information about the true state of the economy.[2] We examine a different assumption that produces an essentially opposite weighting: that the difference between the true state of the economy and each measured statistic is pure *news*, or pure information about the true state of the economy. Under the noise assumption, a statistic with greater variance is given smaller weight because it is assumed to contain more noise; under the news assumption, a statistic with greater variance is given larger weight because it is assumed to contain more news about the true state of the economy.

We play out these contrasting assumptions in our empirical application of choice. The most widely used statistic produced by the U.S. Bureau of Economic Analysis (BEA) is its expenditure-based estimate of the size of the economy, gross domestic product (GDP); however it also produces an income-based estimate of the size of the economy, gross domestic income (GDI), from different information. Computing the value of GDP and GDI would be straigthforward if we could record the value of all the underlying transactions included in the NIPA definition of the size of the economy, and

GDP and GDI would coincide in this case. However all the underlying transactions are not recorded, and measured GDP and measured GDI do not coincide, giving rise to the statistical discrepancy. BEA relies on various surveys, censuses and administrative records, each imperfect, to compute the estimates; differences between the data sources used to produce GDP and GDI, as well as other measurement difficulties, leads to the statistical discrepancy.

Some economists[3] have estimated the growth rate of "true" unobserved GDP as a combination of growth rates of BEA's latest available time series on GDP and GDI, generally concluding that GDI growth should be given more weight than measured GDP growth. Is GDI really the more accurate measure, implying that the emphasis on GDP is misplaced? We argue for caution, as the results in these papers are driven by the noise assumption: the models implicitly assume that since GDP growth has higher variance than GDI growth, it must be noisier, and so should receive a smaller weight. However GDP may have higher variance because it contains more information about "true" unobserved GDP (this is the essence of the news assumption); then measured GDP should receive the higher weight.

We emphasize that, since we never observe the "true" state of the economy, assumptions about news vs. noise are inherently untestable. Indeed, the general version of our model allows the difference between each measured statistic and the "true" state of the economy to be a mixture of news and noise, and we proceed to show that virtually any set of weights can be rationalized by making untestable assumptions about the mixtures. More information must be brought to bear on the problem; otherwise the choice of weights will be totally arbitrary.

We argue that it is possible to gather additional information that may shed some light on whether the differences between "true" GDP and our two measures are mostly news or noise. BEA releases its first national accounts estimates for any given quarter about a month after the quarter closes, and subsequently revises the initial estimates a number of times over the next three to four years (as well as periodically at approximately five year intervals after that).[4] The revisions incorporate more comprehensive and accurate source data, and each revision presumably brings the national accounts estimates closer to the "truth." Following N. Gregory Mankiw and Matthew D. Shapiro (1986), we show that these revisions are in fact mostly news, not noise, and based on this result we make two points. First, when combining the first few releases of GDP and GDI (the estimates that have not been revised many times, and that are of most interest to analysts attempting to gauge the current state of the economy), we should favor the assumptions of the news model, as we know that at least some of the difference between each estimates and "true" GDP is news - the part that will be eliminated later through revisions. Second, we argue that, since the differences between the early release and most current estimates are largely news, the differences between the most current estimates and "true" unobserved GDP are likely news as well. The argument is based on the assumption that observed patterns would continue: if BEA did ultimately acquire exact knowledge of "true" unobserved GDP, we hypothesize that its (hypothetical) ultimate revision from the most current estimates to the truth would be similar to other revisions we have observed in the past. This second point is more speculative than the first, however.

We then proceed to estimate models and compute estimates of "true" unobserved GDP growth under the news and noise assumptions, distinguishing between the first few estimates and the later, more heavily revised estimates that are typically used for historical research. For the first few estimates, the variance of measured GDP exceeds the variance of GDI, so the news assumption dictates that measured GDP receives the higher weight. If we take parameters to be fixed over our full sample, the same holds true for the heavily revised estimates, as the variance of GDP exceeds the variance a GDI, the same result found in prior research on this topic. However the variance of our measured estimates drops dramatically after the early 1980s - see Margaret M. McConnell and

Gabriel Perez-Quiros (2000) - and if we account for this phenomenon, we find that the variance ordering of the two estimates changes: the variance of GDI then exceeds the variance of GDP after the early 1980s. Whether this result will continue going forward remains to be seen, but assuming that it does, when combining the heavily revised later vintages of data in the post-1984 time period, the news assumptions would place more weight on GDI.

Figure 1 plots data from the latest recession and recovery, the 1999-2002 growth rates of the most current BEA data on GDP and GDI, and estimated "true" GDP from the news model.[5] We see that these news model estimates reflect some characteristics of averaging, with the patterns in 2000 in "true" GDP being less erratic than the patterns in either GDP or GDI; indeed, the combined estimates show a smooth downward trend into recession. But we also see in the fourth quarter of 1999, the quarter with the biggest late-cycle boom, "true" GDP growth exceeds the growth rate of either GDP or GDI, and in the third quarter of 2001, the nadir of the recession, "true" GDP growth again takes on a value more extreme than either GDP or GDI. The examples are not anomalies: the variance of our estimated "true" GDP growth exceeds the variance of either GDP growth or GDI growth. This is a natural outcome of the news hypothesis, and since there likely exists other unobserved information about "true" GDP reflected in neither measured GDP or GDI, this estimated variance of "true" GDP represents a lower bound on the actual variance of "true" GDP. If the news hypothesis is true, then, it has clear implications for real business cycle and other economic models that use as an input the variance of GDP.

The rest of the paper is organized as follows. Section I discusses the various models in a simple bivariate setup, drawing out their implications for constructing weighted averages. Section II describes BEA's national income accounts data that we employ in our empirical work, and analyzes whether revisions from the early to later vintages of GDP and GDI are mostly news or noise. Section III reports results from constructing weighted averages of GDP and GDI under both the news model and the noise model. In addition to reporting results from the most current available data, it also reports results from earlier vintages, as these are important and widely followed indicators of the state of the economy. We draw our conclusions in Section IV.

I. Theory: The Competing News and Noise Models

A. Review of News, Noise, and Covariance Assumptions

Let $\Delta y_t^\star$ be the true growth rate of the economy, let Δy_t^k be one of its measured estimates, and let ε_t^k be the difference between the two, so:

$$\Delta y_t^k = \Delta y_t^\star + \varepsilon_t^k.$$

The noise model makes the classical measurement error assumption that $\mathrm{cov}\left(\Delta y_t^\star, \varepsilon_t^k\right) = 0$; this is the precise meaning of the statement that ε_t^k is *noise*. One implication of a noisy estimate Δy_t^k is that it's variance is greater than the variance of the true growth rate of the economy, or $\mathrm{var}\left(\Delta y_t^k\right) > \mathrm{var}\left(\Delta y_t^\star\right)$.

In contrast, if an estimate Δy_t^k were constructed efficiently with respect to a set of information about $\Delta y_t^\star$ (call it $\mathcal{F}_t^k$), then Δy_t^k would be the conditional expectation of $\Delta y_t^\star$ given that information set:

$$\Delta y_t^k \quad = \quad E\left(\Delta y_t^\star | \mathcal{F}_t^k\right).$$

Writing:

$$\Delta y_t^\star \quad = \quad \Delta y_t^k + \zeta_t^k,$$

the term ζ_t^k represents the information about $\Delta y_t^\star$ that is unavailable in the construction of Δy_t^k. Then Δy_t^k and ζ_t^k represent mutually orthogonal pieces of *news* about $\Delta y_t^\star$, employing the terminology in Mankiw and Shapiro (1986), and $\operatorname{cov}\left(\Delta y_t^k, \zeta_t^k\right) = 0$. This leads us to an implication of the news model that we employ later, namely that $\operatorname{cov}\left(\Delta y_t^k, \Delta y_t^\star\right) = \operatorname{var}\left(\Delta y_t^k\right)$. We also have $\operatorname{var}\left(\Delta y_t^\star\right) > \operatorname{var}\left(\Delta y_t^k\right)$, an implication opposite to that of the noise model.

The news model can be written with the notation of the noise model if we take $-\zeta_t^k = \varepsilon_t^k$ and switch this term to the other side of the equation, but the covariance assumption of the noise model will be violated; in fact the error will be perfectly negatively correlated with the missing piece of information about the true growth rate of the economy, so $\operatorname{cov}\left(\Delta y_t^\star, \varepsilon_t^k\right) = \operatorname{cov}\left(\Delta y_t^k + \zeta_t^k, -\zeta_t^k\right) = -\operatorname{var}\left(\varepsilon_t^k\right)$. The variance ordering of the news assumption, $\operatorname{var}\left(\Delta y_t^\star\right) > \operatorname{var}\left(\Delta y_t^k\right)$, will still hold, as:

$$
\begin{aligned}
\operatorname{var}\left(\Delta y_t^k\right) &= \operatorname{var}\left(\Delta y_t^\star\right) + \operatorname{var}\left(\varepsilon_t^k\right) + 2\operatorname{cov}\left(\Delta y_t^\star, \varepsilon_t^k\right) \\
&= \operatorname{var}\left(\Delta y_t^\star\right) - \operatorname{var}\left(\varepsilon_t^k\right).
\end{aligned}
$$

Writing the models in this common notation, and differentiating them by assumptions about the covariance of ε_t^k with $\Delta y_t^\star$, will be useful in discussing the empirical results in the paper.

The pure news and pure noise assumptions are extremes; many intermediate cases could be considered where ε_t^k is part news and part noise, implying differing degrees of negative covariance between $\Delta y_t^\star$ and ε_t^k. We consider a general model that encompasses these intermediate cases in the next subsection.

B. The Mixed News and Noise Model

We consider a model with two estimates of true unobserved GDP, each an efficient estimate plus noise:

$$\begin{aligned}\Delta y_t^1 &= E\left(\Delta y_t^\star | \mathcal{F}_t^1\right) + \varepsilon_t^1, \quad \text{and} \\ \Delta y_t^2 &= E\left(\Delta y_t^\star | \mathcal{F}_t^2\right) + \varepsilon_t^2.\end{aligned}$$

The noise variables ε_t^1 and ε_t^2 are mutually uncorrelated and, naturally, uncorrelated with true unobserved GDP. Taking Δy_t^1 to be GDP and Δy_t^2 to be GDI, the information in $\mathcal{F}_t^1$ likely would consist of personal consumption expenditures, investment, net exports, and the other components that sum to GDP, while the information in $\mathcal{F}_t^2$ likely would consist of wage and salary income, corporate profits, proprietors' income, and the other components that sum to GDI.[6,7] We assume each information set includes a constant, so both Δy_t^1 and Δy_t^2 consistently estimate the mean μ of $\Delta y_t^\star$, and there may be a substantial amount of additional overlap between the two information sets. Consumption growth may be highly correlated with the growth rate of wages and salaries, for example. However a key feature of our model is that it recognizes that the two information sets are not necessarily identical.

To clearly illustrate the main points of the paper, we focus on the simple case where all variables are jointly normally distributed, and where measured GDP and GDI are serially uncorrelated.[8] With normality, the conditional expectation of the true growth rate of the economy is a weighted average of GDP and GDI; netting out means yields:

$$\begin{aligned}E\left(\Delta y_t^\star - \mu | \Delta y_t^1, \Delta y_t^2, \mu\right) &= \\ \widehat{\Delta y_t^\star} - \mu &= \omega_1\left(\Delta y_t^1 - \mu\right) + \omega_2\left(\Delta y_t^2 - \mu\right), \end{aligned} \tag{1}$$

calling the conditional expectation $\widehat{\Delta y_t^\star}$. The weights ω_k can be derived using standard formulas for the population version of ordinary least squares:

$$(2)\quad \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = \begin{pmatrix} \operatorname{var}\left(\Delta y_t^1\right) & \operatorname{cov}\left(\Delta y_t^1, \Delta y_t^2\right) \\ \operatorname{cov}\left(\Delta y_t^1, \Delta y_t^2\right) & \operatorname{var}\left(\Delta y_t^2\right) \end{pmatrix}^{-1} \begin{pmatrix} \operatorname{cov}\left(\Delta y_t^1, \Delta y_t^\star\right) \\ \operatorname{cov}\left(\Delta y_t^2, \Delta y_t^\star\right) \end{pmatrix}$$

$$= \begin{pmatrix} \operatorname{var}\left(\Delta y_t^1\right) & \operatorname{cov}\left(\Delta y_t^1, \Delta y_t^2\right) \\ \operatorname{cov}\left(\Delta y_t^1, \Delta y_t^2\right) & \operatorname{var}\left(\Delta y_t^2\right) \end{pmatrix}^{-1} \begin{pmatrix} \operatorname{var}\left(E\left(\Delta y_t^\star | \mathcal{F}_t^1\right)\right) \\ \operatorname{var}\left(E\left(\Delta y_t^\star | \mathcal{F}_t^2\right)\right) \end{pmatrix},$$

using $\operatorname{cov}\left(\Delta y_t^\star, \varepsilon_t^k\right) = 0$ and the property of efficient estimates that their covariance with the variable they are estimating is simply their variance.

It is useful to introduce some additional notation. Call the covariance between the two estimates σ^2; this arises from the overlap between the information sets used to compute the efficient estimates.[9] The model imposes the condition that the variance of each estimate is at least as large as the covariance between the two; then let $\sigma^2 + \tau_1^2$ and $\sigma^2 + \tau_2^2$ be the variance of the Δy_t^1 and Δy_t^2, respectively. The idiosyncratic variance in each estimate, the τ_k^2 for $k = 1, 2$, arises from two potential sources. The first is the idiosyncratic news in each estimate - the information in each efficient estimate missing from the other. The second source of idiosyncratic variance is the noise, ε_t^k.

Let the fraction of idiosyncratic variance in the kth estimate that is news is be χ_k, so $\chi_k \tau_k^2$ is the variance of idiosyncratic news in Δy_t^k, and $(1 - \chi_k)\,\tau_k^2$ is the variance of measurement error. This χ_k will range from zero, the case where the idiosyncratic variation in the estimate is pure noise, to one, the case where that variation is pure news. Then equation (2) becomes:

$$\begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = \begin{pmatrix} \sigma^2 + \tau_1^2 & \sigma^2 \\ \sigma^2 & \sigma^2 + \tau_2^2 \end{pmatrix}^{-1} \begin{pmatrix} \sigma^2 + \chi_1 \tau_1^2 \\ \sigma^2 + \chi_2 \tau_2^2 \end{pmatrix}.$$

Solving and substituting into (2) gives:

$$\widehat{\Delta y_t^\star} - \mu = \frac{\left(\chi_1 \tau_1^2 + (1-\chi_2)\,\tau_2^2 + \chi_1 \frac{\tau_1^2 \tau_2^2}{\sigma^2}\right)\left(\Delta y_t^1 - \mu\right)}{\tau_1^2 + \tau_2^2 + \frac{\tau_1^2 \tau_2^2}{\sigma^2}} + \frac{\left(\chi_2 \tau_2^2 + (1-\chi_1)\,\tau_1^2 + \chi_2 \frac{\tau_1^2 \tau_2^2}{\sigma^2}\right)\left(\Delta y_t^2 - \mu\right)}{\tau_1^2 + \tau_2^2 + \frac{\tau_1^2 \tau_2^2}{\sigma^2}}. \tag{3}$$

To understand this formula, it is helpful to work through some special cases of interest, which we do next.

It is important to note that not all of the parameters in this model are identified. We observe three moments from the variance-covariance matrix of $[\Delta y_t^1 \quad \Delta y_t^2]$, not enough to pin down the five parameters σ^2, τ_1^2, τ_2^2, χ_1 and χ_2. Imposing values for χ_1 and χ_2 will allow identification of the remaining parameters, and previous attempts to estimate models of this kind have focused on one particular imposition, namely $\chi_1 = \chi_2 = 0$. Then $E\left(\Delta y_t^\star | \mathcal{F}_t^1\right) = E\left(\Delta y_t^\star | \mathcal{F}_t^2\right)$, so the two information sets must coincide, at least in the universe of information that is relevant for predicting $\Delta y_t^\star$. In addition, previous models have assumed that $E\left(\Delta y_t^\star | \mathcal{F}_t^k\right) = \Delta y_t^\star$, for $k = 1, 2$; in this case, the difference between each estimate and the truth, $\Delta y_t^k - \Delta y_t^\star$, is pure noise as in the first example in the previous subsection. We call the general model with this set of assumptions the *pure noise model*, and under this model equation (3) becomes:[10]

$$\widehat{\Delta y_t^\star} - \mu = \frac{\tau_2^2\left(\Delta y_t^1 - \mu\right) + \tau_1^2\left(\Delta y_t^2 - \mu\right)}{\tau_1^2 + \tau_2^2 + \frac{\tau_1^2 \tau_2^2}{\sigma^2}}. \tag{4}$$

In the pure noise model, the weight for one measure is proportional to the idiosyncratic variance of the other measure - since the idiosyncratic variance in each estimate is assumed to be noise, the "noisier" measure is downweighted. The weights on the (net of mean) estimates sum to less than one; as is typical in the classical measurement error model, coefficients on noisy explanatory variables are downweighted. In fact, as the common variance σ^2 approaches zero, the signal-to-noise ratio in the model approaches zero as well, and the formula instructs us to give up on the estimates of GDP and GDI for any given time period, using the overall sample mean as the best estimate for each and every period.

The opposite case is the what we call the *pure news model*, where $\chi_1 = \chi_2 = 1$. The difference between each estimate and the truth, $\Delta y_t^k - \Delta y_t^\star$, is pure news or pure information in this case, as in the second example in the previous subsection. Equation (3) then becomes:

$$\widehat{\Delta y_t^\star} - \mu = \frac{\left(\tau_1^2 + \frac{\tau_1^2\tau_2^2}{\sigma^2}\right)(\Delta y_t^1 - \mu) + \left(\tau_2^2 + \frac{\tau_1^2\tau_2^2}{\sigma^2}\right)(\Delta y_t^2 - \mu)}{\tau_1^2 + \tau_2^2 + \frac{\tau_1^2\tau_2^2}{\sigma^2}}. \tag{5}$$

The weight for each measure is now proportional to its own idiosyncratic variance - the estimate with greater variance contains more news and hence receives a larger weight. And in another result diametrically opposed to that of the noise model, the weights (on the net of mean estimates) sum a number greater than unity. As $\sigma^2 \to 0$ (i.e. as the variance common to the two estimates approaches zero), the weight for each estimate approaches unity. In this case, we are essentially adding together two independent pieces of information about GDP growth. To illustrate, suppose we receive news of a shock that moves $\Delta y_t^\star$ two percent above its mean, and then receive news of another, independent shock that moves $\Delta y_t^\star$ one percent below its mean. The logical estimate of $\Delta y_t^\star$ is then the mean plus one percent - i.e. the sum of the two shocks. In Appendix A we work through another example, of two estimates of GDP growth, each based on the growth rate of a different sector of the economy; if the growth rates of the sectors are uncorrelated, we simply add up the net-of-mean contributions to GDP growth of the two sectors, and then add back in the mean.

Moving back to the more general model, we see by adding μ to equation (3) that in all cases the model instructs us to take a weighted average of the growth rate of GDP, the growth rate of GDI, and μ; the weights on these three variables sum to one. However in some situations we may have little confidence in our estimate of the mean μ, and we may be uncomfortable using it as the third component in our weighted average. One way around this problem is to force the weights on Δy_t^1 and Δy_t^2 to sum to one, with $\omega_2 = 1 - \omega_1$; substituting into (1) and rearranging, we arrive at an expectation that can be computed without knowledge of μ:

$$E\left(\Delta y_t^\star - \Delta y_t^2 | \Delta y_t^1, \Delta y_t^2\right) = \omega_1 \left(\Delta y_t^1 - \Delta y_t^2\right). \tag{6}$$

Adding back in Δy_t^2 yields $\widehat{\Delta y_t^\star}$. The solution to the general model then becomes:

$$\widehat{\Delta y_t^\star} = \frac{\left(\chi_1 \tau_1^2 + (1-\chi_2)\,\tau_2^2\right)\Delta y_t^1 + \left(\chi_2 \tau_2^2 + (1-\chi_1)\,\tau_1^2\right)\Delta y_t^2}{\tau_1^2 + \tau_2^2}. \tag{3'}$$

It turns out that, with the assumptions of the pure noise model, this particular estimator is equivalent to the estimator proposed by Martin R. Weale (1992), who applied it to the case of GDP and GDI the techniques developed in Sir Richard Stone, D.G. Champernowne, and J. E. Meade (1942). Appendix B clarifies the relation between these earlier estimators and those that we propose here.

To close this section, we consider some other special cases of the general model. If $\chi_1 = 1$ and $\chi_2 = 0$, then $\omega_1 = 1$ and $\omega_2 = 0$; placing all the weight on any given estimate amounts to an assumption that the idiosyncratic portion of that estimate is pure news, and the idiosyncratic portion of the other estimate is pure noise. This leads us to a more general statement. Let the ratio of the weights $\frac{\omega_1}{\omega_2} = r$, so:

$$r\left(\chi_1, \chi_2\right) = \frac{\chi_1 \tau_1^2 + (1-\chi_2)\,\tau_2^2 + \chi_1 \frac{\tau_1^2 \tau_2^2}{\sigma^2}}{\chi_2 \tau_2^2 + (1-\chi_1)\,\tau_1^2 + \chi_2 \frac{\tau_1^2 \tau_2^2}{\sigma^2}}, \tag{7}$$

where we've expressed r as a function of χ_1 and χ_2. The following proposition shows that any set of weights can be rationalized by making untestable assumptions about the degree of news and noise in the two measures of the state of the economy:

Proposition 1 *Let r be any non-negative real number, and let $r(\chi_1, \chi_2)$ be given by (7), where τ_1^2, τ_2^2, and σ^2 are each constant, positive real numbers. Then there exists a pair $(\chi_1^\star, \chi_2^\star)$, with $\chi_1^\star \in [0, \quad 1)$ and $\chi_2^\star \in (0, \quad 1]$, such that $r(\chi_1^\star, \chi_2^\star) = r$.*

Proof: We provide an example with a closed form solution. Consider the case where $\chi_2 = 1 - \chi_1$; then $r(\chi_1, \chi_2) = \frac{\chi_1}{1-\chi_1}$. Since $r(\chi_1, \chi_2)$ is a continuous function, $r(0, 1) = 0$, and $\lim_{\chi_1 \to 1} r(\chi_1, 1 - \chi_1) = \infty$, the result holds by theorem 4.23 of Rudin (1953). We have $\chi_1 = \frac{r}{1+r}$, which produces the desired $\chi_1^\star \in [0, \quad 1)$ and $\chi_2^\star \in (0, \quad 1]$ for any non-negative real r.

One set of weights is as justifiable as any other; without further information about the estimates, the choice of weights will be arbitrary. In the empirical work below on GDP and GDI, we do bring further information to bear on the problem, and examine whether the pure news or pure noise model is closer to reality.

II. Data: GDP and GDI

Using expenditure information, GDP is equal to the sum of consumption, investment, government expenditures, and net exports. Using income information, national income is the sum of employee compensation, proprietors' income, rental income, corporate profits and net interest; adding consumption of fixed capital and a few other balancing items to national income produces GDI.[11]

Table 1 summarizes the sequence of vintages of quarterly GDP and GDI data released by BEA. The "advance" estimate of the most current quarter is released about a month after the quarter closes, with the "preliminary" estimate following a month after "advance" and the "final" estimate following a month after "preliminary". Usually in the summer of year $t+1$, all quarters of year t are reopened for the first annual revision;

the second and third annual revisions occur in subsequent years, $t+2$ and $t+3$. Finally, about every five years, all of the accounts data are reopened for benchmark revisions. The benchmarks are a mixture of methodological changes, statistical changes, and the incorporation of previously unavailable data, mainly from the most recent quinquennial economic census. BEA maintains a database of each of the vintages of estimates for GDP, GDI, and various sub-components, extending back to 1978. Our sample extends from this date through 2002, where we stop so we can use third annual revision estimates throughout the sample. Our latest available data series have passed through the 2005 annual revision, and were pulled from the BEA web site in August 2005; figure 2 plots the annualized quarterly growth rates of these latest available nominal GDP (solid line) and nominal GDI (dashed line) numbers.[12] As numerous economists have noted, the growth rate of real GDP became less volatile sometime around 1984, and figure 2 shows that this is evidently the case for nominal GDP and GDI as well. In addition, these nominal data reflect relatively high inflation in the U.S. in the late 1970s and early 1980s.

Before reporting the summary statistics for our data, some additional notation is helpful. Let $\Delta y_t^{1,a}$ be the "advance" estimate of measured GDP growth; $\Delta y_t^{1,p}$ the "preliminary", and so forth, with $\Delta y_t^{1,l}$ being the latest available estimate. We assume that later vintages are better estimates of $\Delta y_t^{\star}$ than are earlier ones, so if we write:

$$\begin{aligned}
\Delta y_t^{1,a} &= \Delta y_t^{\star} + \varepsilon_t^{1,a} \\
\Delta y_t^{1,p} &= \Delta y_t^{\star} + \varepsilon_t^{1,p} \\
&\vdots \\
\Delta y_t^{1,l} &= \Delta y_t^{\star} + \varepsilon_t^{1,l}
\end{aligned}$$

then $\operatorname{var}\left(\varepsilon_t^{1,a}\right) > \operatorname{var}\left(\varepsilon_t^{1,p}\right) > \ldots > \operatorname{var}\left(\varepsilon_t^{1,l}\right)$. Under the pure noise model the ε_t are uncorrelated with $\Delta y_t^\star$, so we should observe the variance of the estimates falling as we move forward in vintage. Under the pure news model the ε_t are highly negatively correlated with $\Delta y_t^\star$, so we should observe the variance of the estimates increasing as we move forward, which would be consistent with the flow of source data BEA uses to compute the estimates. For earlier vintages of estimates, especially the current quarterly ones, BEA lacks data on some components of GDP and GDI. In these cases BEA often uses "trend extrapolations," usually assuming the growth rate for the current quarter is equal to something like the average growth rate over the past five years. Such extrapolations will generally have low variance, not straying too far from the mean, and when BEA receives and substitutes actual data for these extrapolated components in later vintages, the variance of the growth rates will increase. The new source data on some components of GDP and GDI that BEA receives in the later vintages is part of the news in each revision.

Table 2 reports summary statistics on means and variances of growth rates of GDP and GDI, for different vintages (again in percent changes at an annual rate). BEA does not produce "advance" estimates of GDI, as the corporate profits data, and at times the net interest data, are unsatisfactory that close to the end of the quarter. BEA does produce "preliminary" GDI for quarters one to three, but unfortunately not for fourth quarters; for this reason we decline to report GDI results for those vintages.

The results in table 2 are broadly consistent with the news model. Over the full 1978 to 2002 sample, for both GDP and GDI, the variances increase with each move forward in vintage, except when moving from the first annual revision to the second annual. The table also reports results for the 1984Q3 to 2002 sub-sample, excluding the high-variance, high-inflation period in our data.[13] The drop in overall variance in this panel is clearly evident; looking across vintages in this panel, while we do see a drop in the variance of GDP growth as we move from the third annual to the latest available vintage, the evidence still broadly favors the news model for both GDP and GDI. The lower panels of table 2 report means and variances of revisions - the first column reports means and variances of $\Delta y^p - \Delta y^a$, for example. Measured by variance, the largest revisions occur moving from final current quarterly to first annual revision, and in the benchmarks that move from third annual to latest, at least over the full sample. It should be kept in mind that what we call the revision from the third annual to the latest is, for most years, actually the sum of multiple benchmark revisions.

Table 3 follows the presentation of Mankiw and Shapiro (1986) in presenting a correlation matrix of each revision with the full set of vintages, shedding more light on whether revisions to GDP and GDI are news or noise. Under the news model, revisions should be uncorrelated with prior estimates, but should be positively correlated with current and subsequent estimates. So $\Delta y^p - \Delta y^a$ should be correlated with Δy^p, Δy^f, Δy^{ar1}, Δy^{ar2}, Δy^{ar3}, and Δy^l, but not Δy^a; $\Delta y^f - \Delta y^p$ should be correlated with Δy^f, Δy^{ar1}, Δy^{ar2}, Δy^{ar3}, and Δy^l, but not Δy^a or Δy^p, and so on. Under the noise model, the exact opposite is true; hence this correlation table represents a compactly expressed horse race between the two models.

Panels A and B of table 3 show results over the full sample for GDP and GDI; the parenthesis below the correlation estimates in each box now represent t-statistics, not standard errors. We see a large number of statistically significant coefficients in the upper right-hand section in support of the news model, and zero in the lower left-hand section in support of the noise model. Panels C and D report results for the 1984Q3 to 2002 sub-sample; here the evidence in favor of the news model is somewhat less strong, but the data still lean heavily in that direction.

Two points should be kept in mind about this evidence indicating that revisions to GDP and GDI are largely news, not noise. First, under the reasonable assumption that the later vintages are better estimates of $\Delta y_t^\star$ than are the earlier vintages, the evidence indicates that at least part of the difference between each early vintage estimate and "true" unobserved GDP is news, not noise. Based on this evidence, attempts to combine the early vintage estimates such as current quarterly should favor the news model. Second, we think it is also reasonable to draw at least tentative inferences about the differences between "true" unobserved GDP and the *latest available* estimates of GDP and GDI. We've examined seven vintages of GDP and GDI data, finding mostly news through the six revisions as we move from one vintage to the next; it is more probable than not that we'd get news again in a hypothetical ultimate revision from the latest available data to the truth.

III. Estimates of "True" Unobserved GDP

Table 4 reports maximum likelihood estimates of the pure news and pure noise models, for various vintages of data. The likelihood function is the same for both models; only the formulae for the weights to be placed on GDP and GDI differ. We report the four parameter estimates, the weights for each model, and, for the news model, the weights renormalized so they sum to one as in (3'); the renormalized weights for the noise model are just the mirror images of those reported. In addition, we report the variance of the predicted values for "true" GDP growth for each model. For the pure news model, we interpret this quantity as a lower bound on the variance of true GDP growth.[14] Writing:

$$\Delta y_t^{\star} - \mu = \omega_1 \left(\Delta y_t^1 - \mu\right) + \omega_2 \left(\Delta y_t^2 - \mu\right) + \zeta_t, \tag{8}$$

the ζ_t term represents an additional piece of information, the information about $\Delta y_t^{\star}$ contained in neither available estimate. The variance of the ζ_t term is unknown; however we do know that ζ_t is orthogonal to our estimated $\widehat{\Delta y_t^{\star}}$, and so the variance of the actual $\Delta y_t^{\star}$ must be less then the variance of estimated $\widehat{\Delta y_t^{\star}}$. Standard errors in the tables are below each estimate in parentheses, and those for the weights are computed using the

Delta method.

Panel A of table 4 shows results for the full 1978 to 2002 sample. Since our empirical evidence from the previous section favors the news model, where variance is a good rather than a bad, GDP should receive the higher weight for all vintages we examine here, with the weight on GDP especially large for the final current quarterly vintage data. If one did favor the noise model instead, the weights appropriate for the latest available vintage of data are similar to those reported in Weale (1992). These results follow directly from the fact that the variance of GDP exceeds the variance of GDI as reported in table 2.

Panel B reports results from estimation of the models over the 1984Q3 to 2002 sub-sample. This choice of break point was guided by likelihood-ratio tests allowing for breaks in some parameters over potential points in the middle 70% of the sample - see Donald W. K. Andrews (1993). Allowing for a break in σ^2 produces massive increases in the likelihood function, with the greatest increase occuring with a 1984Q3 break for all five vintages we examine. Allowing for a break in μ at the same point as the break in σ^2 (the visual evidence in Figure 2 indicates that these two breaks were roughly coincident), we again see the greatest increase in the likelihood with a 1984Q3 break, again for all five vintages. Evidence for further breaks in τ_1^2 and τ_2^2 was mixed, passing tests of statistical significance for some vintages but not for others; for this reason we report results in table 5 where these idiosyncratic variances were held constant throughout the sample.

A dominant feature here is again the large reduction in common variance σ^2 over the sub-sample; this can be seen most starkly in table 5. The drop in μ due to moderation of inflation is evident as well. More important for the relative weights on GDP and GDI are the changes in the idiosyncratic variances τ_1^2 and τ_2^2. Table 2 showed that, when moving from the full sample to the sub-sample, the variance of GDP drops by more than the variance of GDI, most drastically for the latest available data. For all vintages except the final current quarterly vintage, the variance of GDI becomes larger than the variance of GDP, reversing the full sample ordering; these developments appear in table 4, panel B: the estimated idiosyncatic variance of GDI now exceed the estimated idiosyncratic variance of GDP for these vintages, and the weight on GDI under the *news* model now exceeds the weight on GDI. Table 5 shows that even when we impose the same τ_1^2 and τ_2^2 over the full sample, we get a similar reversal in the relative weights for most vintages when we explicitly model the breaks in μ and σ^2. Indeed, when examining the weights that sum to unity, we see that the full sample results in table 5 are similar to the sub-sample results in table 4, panel B. Only for the final current quarterly vintage do we see no reversals in the weights; the noise nodel consistently favors GDI, and the news model GDP. Since the empirical evidence presented in the previous section in favor of the news model is most relevant for these current quarterly estimates, the evidence for this vintage is most clear cut, although the small size of the idiosyncratic variances makes the standard errors of the weights somewhat large.

Those small idiosyncratic variances in the earlier vintages point to an interesting fact: GDP and GDI become more dissimilar as we move forward in vintage, a feature that is robust to choice of sub-sample. While this may seem to contradict our assumption that both GDP and GDI move closer to the truth as we move forward in vintage, the finding seems consistent with the way BEA constructs the estimates. For the earliest vintage estimates, the data available to BEA is quite limited, and much of it must be used in computing both GDP and GDI, thereby making them quite similar. For example, data on much of services consumption, an expenditure-side component, is missing at the time of the current quarterly estimates, so BEA borrows data from the income-side, using employment, hours and earnings as a substitute for many sub-components of services. In later vintages, when more complete and appropriate data on components of GDP and GDI become available, the data overlap between the two measures becomes smaller, and each measure becomes a better estimate of true unobserved GDP even as they become less similar.

One of the most interesting aspects of these results pertains to the variance of the "true" GDP growth under the differing sets of assumptions, especially for the latest available data. In the post 1984 period, we see that var $\left(\widehat{\Delta y_t^\star}\right)$ from the news model is substantially larger than σ^2 and substantially larger than the variance of GDP growth, often employed to estimate the variance of the growth rate of the economy in prior research. This is not surprising, as the news model weights more heavily the component series with higher variance and uses weights that sum to more than one. If the news model is true, this relatively large variance represents a lower bound on the variance of "true" GDP growth, a fact with potentially important implications for a wide class of economic models that depend importantly on the variance of the growth rate of the economy, for example real business cycle models.

IV. Conclusions

We have developed a new model for estimating aggregate economic activity - what we have called "true" unobserved GDP - as a weighted average of measured GDP and another measure of aggregate economic activity, GDI. In principle, GDP and GDI should coincide, but because of differences in data flow they do not. GDP and GDI rely on different source data that are not always completely compatible, giving rise to what is called the statistical discrepancy. Combining GDP and GDI in some way may produce an estimate that is superior to either one in isolation, but previous attempts to do so have made the strong implicit assumption that the difference between "true" GDP and each measured statistic is pure noise, or completely uncorrelated with "true" GDP. The model developed in this paper allows for the possibility that the difference between the "true" GDP and each measured statistic is partly or pure news, or correlated with "true" GDP. If this is true, then our model would weight more heavily the statistic with higher variance, as it is assumed to contain more news about "true" GDP, in contrast to previous models, which always weight less heavily the statistic with higher variance, as it is assumed to contain more measurement error.

We provide evidence that BEA's long sequence of revisions to GDP and GDI are largely news, showing that at least part of differences between "true" GDP and the early-release, unrevised estimates of GDP and GDI are news; we further argue on the basis of continuity that the differences between "true" GDP and the most current vintages of GDP and GDI are likely news as well. However this evidence is certainly not definitive - since "true" GDP is unobserved, we can at most make educated guesses about it, and some parameters of the news and noise models we estimate will remain unidentified. As such, some type of Bayesian combining of the different models may be a promising way to proceed in future research, or Minimax estimation over the unidentified parameters of the models, incorporating the evidence on revisions presented in this paper into prior distributions.[15]

Our results have obvious uses for analysts of the current state of the economy and business cycles, and have broader implications for many economic models. If the news hypothesis is true, we show that the true variance of the growth rate of the economy is not equal to the variance of measured GDP growth, as is often assumed for example in real business cycle models; the true variance is actually higher.

While our empirical results focus on GDP, some type of news model applies more generally whenever the goal is to combine the information in multiple, efficiently-constructed estimates of a variable, each based on incomplete and non-identical information. To further emphasize the generality of the intuition behind our results, consider the well known index of coincident indicators constructed by James H. Stock and Mark W. Watson (1989), discussed in James D. Hamilton (1994) and Francis X. Diebold and Glenn D. Rudebusch (1996). Stock and Watson decompose each of four time series into a common factor plus an idiosyncratic component; a time series that covaries relatively less with the other three will receive less weight in the common factor and have higher idiosyncratic variance. Stock and Watson define the state of the economy as this common factor, so a series with greater (relative) idiosyncratic variance receives less weight in this construct. Is this best weighting? There may be good reasons to define the state of the economy as this common factor, following the venerable tradition of Arthur F. Burns and Wesley C. Mitchell (1946). However if we are willing to define the state of the economy as something other than this common factor, the answer to this question is unclear: if the differences between the true state of the economy and these series are noise, the Stock and Watson approach is appropriate, but if these differences are news, then the variables that contain much idiosyncratic variation are uniquely informative about the state of the economy, and should be weighted *more* heavily.

The examples discussed in this paper illustrate that the noise assumption is often implicit in models of imperfect measurement (in state space models often entering through the assumed orthogonality of the errors of the observation equations with the errors of the state equations); a contribution of this paper is to pull this hidden assumption out into the open, so that economists and statisticians can thoroughly assess its validity. Seemingly innocuous econometric assumptions can imply that the difference between truth and measurement is noise; econometric estimators generally treat variance as a bad, and the noise assumption does as well. We have identified circumstances where this assumption is inappropriate, where variance should be treated as a good, and have taken some steps towards deriving estimators appropriate for these situations.

References

[1] Andrews, Donald W.K. “Tests for Parameter Instability and Structural Change with Unknown Change Point” *Econometrica*, 1993 (61), pp. 821-856.

[2] Burns, A. F., and W. C. Mitchell. Measuring Business Cycles. New York, NBER (1946).

[3] Byron, Ray. “The Estimation of Large Social Accounts Matrices.” *Journal of the Royal Statistical Society*, series A (1978), vol. 141, part 3, pp. 359-367.

[4] Diebold, Francis X. and Rudebusch, Glenn D. “Measuring Business Cycles: A Modern Perspective.” *Review of Economics and Statistics*, 1996 (101), pp. 67-77.

[5] Dynan, Karen E. and Elmendorf, Douglas W. “Do Provisional Estimates of Output Miss Economic Turning Points?” Federal Reserve Board of Governors, working paper, 2001.

[6] Fixler, Dennis J. and Grimm, Bruce T. “Reliability of GDP and Related NIPA Estimates” *Survey of Current Business*, January 2002, pp. 9-27.

[7] Fixler, Dennis J. and Grimm, Bruce T. “Revisions, Rationality, and Turning Points in GDP.” BEA working paper 2003-01, 2003.

[8] Hamilton, James D. Time Series Analysis. Princeton, Princeton University Press (1994).

[9] Harvey, Andrew C. Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge, Cambridge University Press (1989).

[10] Howrey, E. Philip. “The Accuracy of the Government’s Estimates of GDP.” University of Michigan, working paper, December 2003.

[11] Lehmann, E. L., and Casella, George. Theory of Point Estimation. New York, Springer (1998).

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

[13] Mankiw, N. Gregory and Shapiro, Matthew D. "News or Noise: An Analysis of GNP Revisions" *Survey of Current Business*, May 1986, pp. 20-25.

[14] McConnell, Margaret M., and Perez-Quiros, Gabriel. "Output Fluctuations in the United States: What Has Changed Since the Early 1980s?" *American Economic Review*, 2000 (90), pp. 1464-1476.

[15] Rudin, Walter. Principles of Mathematical Analysis. New York, McGraw-Hill (1953).

[16] Smith, Richard J.; Weale, Martin R.; and Satchell, Steven E. "Measurement Error with Accounting Constraints: Point and Interval Estimation for Latent Data with an Application to U.K. Gross Domestic Product" *Review of Economic Studies*, 1998 (65), pp. 109-134.

[17] Stock, James H. and Watson, Mark W. "New Indexes of Coincident anf Leading Economic Indicators." In O. Blanchard and S. Fischer (eds.), NBER Macroeconomics Annual (Cambridge, MA: MIT Press, 1989), pp. 351-394.

[18] Stone, Richard. "Nobel Memorial Lecture 1984: The Accounts of Society." *Journal of Applied Econometrics*, 1986 (1), pp. 5-28.

[19] Stone, Richard; Champernowne, D. G.; and Meade, J. E. "The Precision of National Income Estimates" *Review of Economic Studies*, 1942 (9), pp. 111-125.

[20] Watson, Mark W. "Uncertainty in Model-Based Seasonal Adjustment Procedures and Construction of Minimax Filters." *Journal of the American Statistical Association*, (1987), vol. 82, Applications, pp. 395-408.

[21] Weale, Martin. "Testing Linear Hypotheses on National Accounts Data." *Review of Economics and Statistics*, 1985 (90), pp. 685-689.

[22] Weale, Martin. "Estimation of Data Measured With Error and Subject to Linear Restrictions." *Journal of Applied Econometrics*, 1992 (7), pp. 167-174.

Appendix A: A Simple Example of the Bivariate News Model

We will consider two efficient estimates of true GDP growth, one based on consumption growth, and the other based on the growth rate of investment. After constructing each efficient estimate, we will discuss how to produce the improved estimate of true GDP growth by combining them with equation (5).

Let ΔC_t, ΔI_t, ΔG_t, and ΔNX_t be the contributions to true GDP growth $\Delta y_t^\star$ of consumption, investment, government, and net exports, so:

$$\Delta y_t^\star = \Delta C_t + \Delta I_t + \Delta G_t + \Delta NX_t.$$

Our first efficient estimate of $y_t^\star$, Δy_t^1, is based on $\mathcal{F}_t^1 = [1, \quad \Delta C_t]$, a constant and consumption growth, and the second is based on $\mathcal{F}_t^2 = [1, \quad \Delta I_t]$, a constant and investment growth; the constant in either information set reveals μ, the mean of $y_t^\star$, as well as the means of the component growth rates. Then our efficient estimates will take the form:

$$\begin{aligned}
\Delta y_t^1 &= \mu + (\Delta C_t - \mu_C) + E\left(\Delta I_t - \mu_I | \mathcal{F}_t^1\right) + E\left(\Delta G_t + \Delta NX_t - \mu_G - \mu_{NX} | \mathcal{F}_t^1\right); \\
\Delta y_t^2 &= \mu + (\Delta I_t - \mu_I) + E\left(\Delta C_t - \mu_C | \mathcal{F}_t^2\right) + E\left(\Delta G_t + \Delta NX_t - \mu_G - \mu_{NX} | \mathcal{F}_t^2\right).
\end{aligned}$$

For simplicity, we will examine the case where neither $\mathcal{F}_t^1$ nor $\mathcal{F}_t^2$ contains any useful information about $\Delta G_t + \Delta NX_t - \mu_G - \mu_{NX}$, so the last term in each of the above expressions is zero, and $\Delta G_t + \Delta NX_t - \mu_G - \mu_{NX}$ represents the information about $y_t^\star$ contained in neither of our two estimates.

The relation between ΔC_t and ΔI_t determines the nature of the efficient estimates and weights on Δy_t^1 and Δy_t^2 in equation (5). Consider first the case where these variables are independent. Then:

$$\begin{aligned}\Delta y_t^1 &= \mu + (\Delta C_t - \mu_C) \quad \text{and:} \\ \Delta y_t^2 &= \mu + (\Delta I_t - \mu_I)\,.\end{aligned}$$

There is no information common to $\mathcal{F}_t^1$ and $\mathcal{F}_t^2$, no covariance between the estimates, so $\sigma^2 = 0$. Equation (5) instructs us to remove the mean from each estimate, and then simply add them. Adding back in the mean, we have the natural result:

$$\widehat{\Delta y_t^\star} = \mu + (\Delta C_t - \mu_C) + (\Delta I_t - \mu_I)\,.$$

The weight on each estimate (net of mean) is just one; as mentioned in the previous subsection, this is the case where we are essentially adding independent contributions to GDP growth.

Next consider the case where ΔC_t and ΔI_t are perfectly correlated, so:

$$(\Delta I_t - \mu_I) = a\,(\Delta C_t - \mu_C)\,,$$

where a is some constant. Then:

$$\begin{aligned}\Delta y_t^1 &= \mu + (1+a)\,(\Delta C_t - \mu_C) &= \mu + (\Delta C_t - \mu_C) + (\Delta I_t - \mu_I) \quad \text{and:} \\ \Delta y_t^2 &= \mu + (1+\tfrac{1}{a})\,(\Delta I_t - \mu_I) &= \mu + (\Delta C_t - \mu_C) + (\Delta I_t - \mu_I)\,.\end{aligned}$$

Given that $\Delta y_t^1 = \Delta y_t^2$, taking a weighted average of the two produces the same estimate as long as the weights in the average sum to one. There is no idiosyncratic variance to either estimate, so $\tau_1^2 = \tau_2^2 = 0$, and equation (5) instructs us to use a weight of 0.5 for each estimate.[16]

Finally consider the general linear case. In this case:

$$\begin{aligned} E\left(\Delta I_t - \mu_I | \mathcal{F}_t^1\right) &= a\left(\Delta C_t - \mu_C\right) \quad \text{and:} \\ E\left(\Delta C_t - \mu_C | \mathcal{F}_t^2\right) &= b\left(\Delta I_t - \mu_I\right) \end{aligned}$$

Least squares projections tell us that $a = \frac{\sigma_{ci}}{\sigma_c^2}$, where σ_{ci} is the covariance between ΔI_t and ΔC_t, and σ_c^2 is the variance of ΔC_t. Similarly, $b = \frac{\sigma_{ci}}{\sigma_i^2}$, where σ_i^2 is the variance of ΔI_t, and the fraction of the variance of each variable explained by the other, R^2, is $\frac{\sigma_{ci}^2}{\sigma_i^2 \sigma_c^2}$. The efficient estimates of $\Delta y_t^\star$ are:

$$\begin{aligned} \Delta y_t^1 &= \mu + (1+a)\left(\Delta C_t - \mu_C\right) \quad \text{and:} \\ \Delta y_t^2 &= \mu + (1+b)\left(\Delta I_t - \mu_I\right). \end{aligned}$$

The variance parameters of the news model are identified from the following relations:

$$\begin{aligned} \sigma^2 &= \operatorname{cov}\left(\Delta y_t^1, \Delta y_t^2\right) = (1+a)(1+b)\sigma_{ci}, \\ \tau_1^2 &= \operatorname{var}\left(\Delta y_t^1\right) - \operatorname{cov}\left(\Delta y_t^1, \Delta y_t^2\right) = (1+a)^2\sigma_c^2 - (1+a)(1+b)\sigma_{ci} \quad \text{and:} \\ \tau_2^2 &= \operatorname{var}\left(\Delta y_t^2\right) - \operatorname{cov}\left(\Delta y_t^1, \Delta y_t^2\right) = (1+b)^2\sigma_i^2 - (1+a)(1+b)\sigma_{ci}. \end{aligned}$$

Substituting $a = \frac{\sigma_{ci}}{\sigma_c^2}$ and $b = \frac{\sigma_{ci}}{\sigma_i^2}$, we see that both $\tau_1^2 > 0$ and $\tau_2^2 > 0$ if $\sigma_{ci}^2 < \sigma_i^2\sigma_c^2$, or if $R^2 < 1$. If $R^2 = 1$, we are clearly back to the perfect correlation case with $\tau_1^2 = 0$ and $\tau_2^2 = 0$; if $R^2 = 0$, we are back to independence with $\sigma^2 = 0$. In all intermediate cases, the sum of the two weights (net of mean) will range between 1 and 2.

It should be pointed out that, when combining Δy_t^1 and Δy_t^2 in this particular example, using equation (5) is not the most natural way to proceed. An easier and more intuitive procedure would be to set $a\left(\Delta C_t - \mu_C\right)$ to zero in Δy_t^1, set $b\left(\Delta I_t - \mu_I\right)$ to zero in Δy_t^2, and then combine, producing:

$$\widehat{\Delta y_t^\star} = \mu + (\Delta C_t - \mu_C) + (\Delta I_t - \mu_I)\,.$$

This is the best possible estimate of $\widehat{\Delta y_t^\star}$ given the information in $\mathcal{F}_t^1$ and $\mathcal{F}_t^2$, so any estimate based on (5) can only be worse. This result highlights one of the key assumptions of the model: it assumes that the econometrician does not have enough information to set to zero or re-weight individual components of either estimate Δy_t^k; the econometrician must take each Δy_t^k in its totality. Considering different weights for different components of GDP and GDI is another interesting avenue for future research.

Appendix B: Relation to Earlier Work Based on Stone, Champernowne, and Meade (1942)

Equation (3') with the pure noise assumptions yields $\widehat{\Delta y_t^\star} = \frac{\tau_2^2 \Delta y_t^1 + \tau_1^2 \Delta y_t^2}{\tau_1^2 + \tau_2^2}$, essentially the estimator presented in Weale (1992).[17] This paper applied to the case of U.S. GDP and GDI the techniques developed in Stone, Champernowne, and Meade (1942) and Byron (1978); see also Weale (1985), and Smith, Satchell, and Weale (1998). In the general case, Stone et al (1942) considered a row vector of estimates x that should but do not satisfy the set of accounting constraints $Ax = 0$. They produce a new set of estimates $\widetilde{x^\star}$ that satisfy the constraints by solving the constrained quadratic minimization problem:

$$\text{(B1)} \qquad \underset{\widetilde{x^\star}}{\text{MIN}} \quad \left(\widetilde{x^\star} - x\right)' V^{-1} \left(\widetilde{x^\star} - x\right) \qquad \text{S.T.} \quad A\widetilde{x^\star} = 0.$$

The matrix V represents a variance-covariance matrix of $x^\star - x$, where $x^\star$ is the vector of "true" values estimated by x, so V^{-1} is an estimate of "precision". The case at hand maps to this framework with the minimization problem looking like:

$$\underset{\widetilde{\Delta y_t^{1\star}}, \widetilde{\Delta y_t^{2\star}}}{\text{MIN}} \quad \begin{pmatrix} \widetilde{\Delta y_t^{1\star}} - \Delta y_t^1 & \widetilde{\Delta y_t^{2\star}} - \Delta y_t^2 \end{pmatrix} V^{-1} \begin{pmatrix} \widetilde{\Delta y_t^{1\star}} - \Delta y_t^1 \\ \widetilde{\Delta y_t^{2\star}} - \Delta y_t^2 \end{pmatrix}$$
$$\text{S.T.} \qquad \widetilde{\Delta y_t^{1\star}} - \widetilde{\Delta y_t^{2\star}} = 0.$$

Substituting the constraint into the objective function, we have:

$$\text{(B2)} \qquad \underset{\widetilde{\Delta y_t^\star}}{\text{MIN}} \quad \begin{pmatrix} \widetilde{\Delta y_t^\star} - \Delta y_t^1 & \widetilde{\Delta y_t^\star} - \Delta y_t^2 \end{pmatrix} V^{-1} \begin{pmatrix} \widetilde{\Delta y_t^\star} - \Delta y_t^1 \\ \widetilde{\Delta y_t^\star} - \Delta y_t^2 \end{pmatrix},$$

with $\widetilde{\Delta y_t^\star} = \widetilde{\Delta y_t^{1\star}} = \widetilde{\Delta y_t^{2\star}}$. The judgement in this approach involves the choice of V. Stone et al (1942) are not so specific in their recommendations, but it seems logical to use estimates of the variance of measurement errors, as defined in the noise model, to compute V, and this is the tack taken by much of the literature following Stone et al (1942). The main point of this paper is that it is also important to consider the relative information content of the different estimates: if one estimate contains much more news than the other estimate, we may want to adjust that estimate less than the other, even if it contains more noise as well. Weale (1992) assumes the idiosyncratic variances of GDP and GDI, the τ_k^2, are measurement errors, as in the noise model above. Under these assumptions, we have:

$$V = \begin{pmatrix} \tau_1^2 & 0 \\ 0 & \tau_2^2 \end{pmatrix}.$$

Solving the quadratic minimization problem with this V, we have $\widetilde{\Delta y_t^\star} = \frac{\tau_2^2 \Delta y_t^1 + \tau_1^2 \Delta y_t^2}{\tau_1^2 + \tau_2^2}$, the same result as the restricted pure noise model.

Problem (B2) is a different minimization problem than the least squares minimization problems that we solve in this paper, where we solve for the weights in (1) or (6) and then compute the predicted values $\widehat{\Delta y_t^\star}$; (B2) solves for $\widetilde{\Delta y_t^\star}$ directly, leaving the weights implicit. In solving for the weights in (1) or (6), assumptions must be made about the covariances between $\Delta y_t^\star$ and the estimates Δy_t^k, whereas in (B2) assumptions must be made about V; as we have seen, when these assumptions are equivalent and when some constraints are applied to (1), the two approaches can give the same result. Comparing the Stone, Champernowne, and Meade (1942) approach with the approach taken here, in a more general setting such as in (B1), is beyond the scope of this paper, but would be an interesting avenue for future research.

Table 1: Summary of Vintages

Vintage	Variable Name
Advance Current Quarterly	Δy^{a}
Preliminary Current Quarterly	Δy^{p}
Final Current Quarterly	Δy^{f}
First Annual Revision	Δy^{ar1}
Second Annual Revision	Δy^{ar2}
Third Annual Revision	Δy^{ar3}
Latest Available	Δy^{l}

Table 2: Summary Statistics,
Growth Rates of GDP and GDI, 1978-2002

Measure		Δy^a	Δy^p	Δy^f	Δy^{ar1}	Δy^{ar2}	Δy^{ar3}	Δy^l
GDP	mean	6.22	6.41	6.46	6.50	6.54	6.63	6.74
	variance	11.55	12.73	13.19	14.38	14.21	14.88	16.18
GDI	mean			6.55	6.60	6.67	6.66	6.75
	variance			12.60	13.91	13.59	14.07	15.86

Growth Rates of GDP and GDI, 1984Q3-2002

Measure		Δy^a	Δy^p	Δy^f	Δy^{ar1}	Δy^{ar2}	Δy^{ar3}	Δy^l
GDP	mean	5.17	5.34	5.32	5.37	5.42	5.47	5.56
	variance	3.62	4.07	4.24	4.40	4.24	4.50	4.31
GDI	mean			5.48	5.48	5.56	5.54	5.58
	variance			3.92	4.48	4.51	5.12	5.51

Revisions from Previous Vintage, 1978-2002

Measure		Δy^p	Δy^f	Δy^{ar1}	Δy^{ar2}	Δy^{ar3}	Δy^l
GDP	mean	0.20	0.04	0.04	0.05	0.08	0.11
	variance	0.69	0.16	1.35	0.72	0.52	1.56
GDI	mean			0.05	0.07	-0.01	0.09
	variance			1.56	1.07	0.98	2.04

Revisions from Previous Vintage, 1984Q3-2002

Measure		Δy^p	Δy^f	Δy^{ar1}	Δy^{ar2}	Δy^{ar3}	Δy^l
GDP	mean	0.17	-0.02	0.05	0.05	0.05	0.08
	variance	0.43	0.11	1.06	0.59	0.46	0.71
GDI	mean			0.00	0.08	-0.03	0.05
	variance			1.27	1.00	0.89	0.78

Table 3: Correlations between Growth Rates and Revisions
Panel A: GDP, 1978-2002

Revision	Δy^a	Δy^p	Δy^f	Δy^{ar1}	Δy^{ar2}	Δy^{ar3}	Δy^l
$\Delta y^p - \Delta y^a$	**0.09**	**0.32**	**0.31**	**0.30**	**0.34**	**0.32**	**0.27**
	(0.87)	(3.30)	(3.27)	(3.09)	(3.56)	(3.36)	(2.75)
$\Delta y^f - \Delta y^p$	**0.11**	**0.11**	**0.22**	**0.19**	**0.18**	**0.14**	**0.08**
	(1.07)	(1.07)	(2.18)	(1.89)	(1.80)	(1.42)	(0.79)
$\Delta y^{ar1} - \Delta y^f$	**-0.01**	**-0.01**	**-0.02**	**0.29**	**0.26**	**0.24**	**0.13**
	(-0.11)	(-0.12)	(-0.18)	(2.98)	(2.66)	(2.43)	(1.31)
$\Delta y^{ar2} - \Delta y^{ar1}$	**-0.15**	**-0.10**	**-0.10**	**-0.14**	**0.08**	**0.09**	**0.05**
	(-1.46)	(-0.98)	(-1.02)	(-1.39)	(0.84)	(0.87)	(0.52)
$\Delta y^{ar3} - \Delta y^{ar2}$	**0.09**	**0.07**	**0.05**	**0.02**	**0.03**	**0.21**	**0.20**
	(0.85)	(0.70)	(0.49)	(0.23)	(0.28)	(2.17)	(2.07)
$\Delta y^l - \Delta y^{ar3}$	**0.16**	**0.12**	**0.10**	**-0.00**	**-0.03**	**-0.03**	**0.28**
	(1.59)	(1.20)	(0.97)	(-0.03)	(-0.26)	(-0.26)	(2.94)

Panel B: GDI, 1978-2002

Revision	Δy^f	Δy^{ar1}	Δy^{ar2}	Δy^{ar3}	Δy^l
$\Delta y^{ar1} - \Delta y^f$	**-0.03**	**0.31**	**0.26**	**0.25**	**0.16**
	(-0.28)	(3.21)	(2.72)	(2.51)	(1.64)
$\Delta y^{ar2} - \Delta y^{ar1}$	**-0.13**	**-0.18**	**0.10**	**0.07**	**0.05**
	(-1.31)	(-1.82)	(0.98)	(0.73)	(0.48)
$\Delta y^{ar3} - \Delta y^{ar2}$	**-0.03**	**-0.04**	**-0.07**	**0.20**	**0.07**
	(-0.26)	(-0.43)	(-0.67)	(1.99)	(0.65)
$\Delta y^l - \Delta y^{ar3}$	**0.15**	**0.08**	**0.07**	**-0.02**	**0.34**
	(1.52)	(0.81)	(0.65)	(-0.24)	(3.53)

Table 3: Correlations between Growth Rates and Revisions
Panel C: GDP, 1984Q3-2002

Revision	Δy^a	Δy^p	Δy^f	Δy^{ar1}	Δy^{ar2}	Δy^{ar3}	Δy^l
$\Delta y^p - \Delta y^a$	**0.01**	**0.33**	**0.37**	**0.35**	**0.35**	**0.28**	**0.23**
	(0.07)	(3.00)	(3.37)	(3.19)	(3.21)	(2.49)	(2.04)
$\Delta y^f - \Delta y^p$	**-0.04**	**0.04**	**0.21**	**0.14**	**0.13**	**0.07**	**-0.01**
	(-0.37)	(0.38)	(1.79)	(1.22)	(1.08)	(0.60)	(-0.12)
$\Delta y^{ar1} - \Delta y^f$	**-0.20**	**-0.20**	**-0.21**	**0.28**	**0.23**	**0.18**	**0.16**
	(-1.76)	(-1.71)	(-1.86)	(2.49)	(2.00)	(1.60)	(1.35)
$\Delta y^{ar2} - \Delta y^{ar1}$	**-0.16**	**-0.16**	**-0.16**	**-0.23**	**0.14**	**0.18**	**0.14**
	(-1.39)	(-1.34)	(-1.38)	(-2.04)	(1.17)	(1.53)	(1.17)
$\Delta y^{ar3} - \Delta y^{ar2}$	**0.03**	**-0.04**	**-0.06**	**-0.12**	**-0.07**	**0.25**	**0.19**
	(0.24)	(-0.31)	(-0.53)	(-1.02)	(-0.58)	(2.21)	(1.63)
$\Delta y^l - \Delta y^{ar3}$	**-0.05**	**-0.09**	**-0.12**	**-0.16**	**-0.20**	**-0.25**	**0.15**
	(-0.43)	(-0.78)	(-1.06)	(-1.37)	(-1.77)	(-2.21)	(1.28)

Panel D: GDI, 1984Q3-2002

Revision	Δy^f	Δy^{ar1}	Δy^{ar2}	Δy^{ar3}	Δy^l
$\Delta y^{ar1} - \Delta y^f$	**-0.16**	**0.38**	**0.27**	**0.20**	**0.14**
	(-1.38)	(3.51)	(2.42)	(1.77)	(1.22)
$\Delta y^{ar2} - \Delta y^{ar1}$	**-0.12**	**-0.23**	**0.24**	**0.23**	**0.20**
	(-0.99)	(-1.99)	(2.14)	(1.97)	(1.71)
$\Delta y^{ar3} - \Delta y^{ar2}$	**0.00**	**-0.07**	**-0.07**	**0.35**	**0.26**
	(0.02)	(-0.56)	(-0.58)	(3.19)	(2.26)
$\Delta y^l - \Delta y^{ar3}$	**0.10**	**0.02**	**-0.01**	**-0.10**	**0.28**
	(0.89)	(0.17)	(-0.05)	(-0.82)	(2.49)

Table 4: Estimates of True Unobserved GDP Growth,
Panel A: 1978-2002

					Noise Model			News Model				
Vintage	μ	σ^2	τ_1^2	τ_2^2	w_{GDP}	w_{GDI}	var $\widehat{\Delta y^\star}$	w_{GDP}	w_{GDI}	var $\widehat{\Delta y^\star}$	w_{GDP}^{sum1}	w_{GDI}^{sum1}
Final Curr. Qtrly.	**6.53**	**12.42**	**0.74**	**0.12**	**0.14**	**0.85**	**12.37**	**0.86**	**0.15**	**13.22**	**0.86**	**0.14**
	(0.36)	(1.80)	(0.35)	(0.33)	(0.38)	(0.39)		(0.38)	(0.39)		(0.38)	(0.38)
First Annual	**6.57**	**13.52**	**0.82**	**0.34**	**0.29**	**0.70**	**13.33**	**0.71**	**0.30**	**14.49**	**0.71**	**0.29**
	(0.37)	(1.97)	(0.42)	(0.41)	(0.34)	(0.35)		(0.34)	(0.35)		(0.35)	(0.35)
Second Annual	**6.63**	**13.06**	**1.06**	**0.45**	**0.29**	**0.68**	**12.82**	**0.71**	**0.32**	**14.35**	**0.70**	**0.30**
	(0.37)	(1.91)	(0.48)	(0.46)	(0.29)	(0.30)		(0.29)	(0.30)		(0.30)	(0.30)
Third Annual	**6.66**	**13.25**	**1.60**	**0.79**	**0.32**	**0.64**	**12.77**	**0.68**	**0.36**	**15.16**	**0.67**	**0.33**
	(0.37)	(1.98)	(0.62)	(0.59)	(0.23)	(0.24)		(0.23)	(0.24)		(0.24)	(0.24)
Latest Available	**6.74**	**14.00**	**2.13**	**1.81**	**0.43**	**0.51**	**13.12**	**0.57**	**0.49**	**17.07**	**0.54**	**0.46**
	(0.39)	(2.14)	(0.83)	(0.81)	(0.18)	(0.19)		(0.18)	(0.19)		(0.20)	(0.20)

Panel B: 1984Q3-2002

					Noise Model			News Model				
Vintage	μ	σ^2	τ_1^2	τ_2^2	w_{GDP}	w_{GDI}	var $\widehat{\Delta y^\star}$	w_{GDP}	w_{GDI}	var $\widehat{\Delta y^\star}$	w_{GDP}^{sum1}	w_{GDI}^{sum1}
Final Curr. Qtrly.	**5.43**	**3.65**	**0.55**	**0.23**	**0.28**	**0.68**	**3.54**	**0.72**	**0.32**	**4.32**	**0.71**	**0.29**
	(0.23)	(0.63)	(0.22)	(0.21)	(0.24)	(0.27)		(0.24)	(0.27)		(0.26)	(0.26)
First Annual	**5.43**	**3.70**	**0.64**	**0.71**	**0.48**	**0.43**	**3.44**	**0.52**	**0.57**	**4.81**	**0.47**	**0.53**
	(0.23)	(0.67)	(0.29)	(0.30)	(0.19)	(0.18)		(0.19)	(0.18)		(0.20)	(0.20)
Second Annual	**5.48**	**3.45**	**0.75**	**1.02**	**0.51**	**0.37**	**3.10**	**0.49**	**0.63**	**4.88**	**0.42**	**0.58**
	(0.23)	(0.65)	(0.33)	(0.35)	(0.16)	(0.15)		(0.16)	(0.15)		(0.17)	(0.17)
Third Annual	**5.50**	**3.36**	**1.08**	**1.70**	**0.51**	**0.33**	**2.85**	**0.49**	**0.67**	**5.66**	**0.39**	**0.61**
	(0.23)	(0.68)	(0.43)	(0.48)	(0.13)	(0.11)		(0.13)	(0.11)		(0.14)	(0.14)
Latest Available	**5.57**	**3.27**	**0.98**	**2.16**	**0.57**	**0.26**	**2.75**	**0.43**	**0.74**	**5.94**	**0.31**	**0.69**
	(0.23)	(0.67)	(0.44)	(0.54)	(0.14)	(0.10)		(0.14)	(0.10)		(0.13)	(0.13)

Table 5: Static Estimates of True Unobserved GDP Growth 1984Q3 Break in μ and σ^2

							News Model	
Vintage	μ(pre84Q3)	μ(post84Q3)	σ^2(pre84Q3)	σ^2(post84Q3)	τ_1^2	τ_2^2	w_{GDP}^{sum1}	w_{GDI}^{sum1}
Final Curr. Qtrly.	**9.62**	**5.43**	**24.74**	**3.62**	**0.59**	**0.28**	**0.68**	**0.32**
	(0.98)	(0.23)	(7.13)	(0.63)	(0.23)	(0.22)	(0.25)	(0.25)
First Annual	**9.74**	**5.43**	**28.41**	**3.75**	**0.59**	**0.58**	**0.50**	**0.50**
	(1.05)	(0.23)	(8.31)	(0.67)	(0.25)	(0.25)	(0.20)	(0.20)
Second Annual	**9.77**	**5.48**	**27.63**	**3.49**	**0.70**	**0.82**	**0.46**	**0.54**
	(1.04)	(0.23)	(8.10)	(0.64)	(0.28)	(0.29)	(0.17)	(0.17)
Third Annual	**9.89**	**5.50**	**27.74**	**3.44**	**1.02**	**1.36**	**0.43**	**0.57**
	(1.04)	(0.23)	(8.17)	(0.67)	(0.36)	(0.38)	(0.14)	(0.14)
Latest Available	**10.08**	**5.57**	**30.23**	**3.06**	**1.40**	**2.54**	**0.35**	**0.65**
	(1.09)	(0.23)	(8.61)	(0.67)	(0.53)	(0.61)	(0.13)	(0.13)

	1978Q1-1984Q2						1984Q3-2002Q4					
	Noise Model			News Model			Noise Model			News Model		
Vintage	w_{GDP}	w_{GDI}	var $\widehat{\Delta y^\star}$	w_{GDP}	w_{GDI}	var $\widehat{\Delta y^\star}$	w_{GDP}	w_{GDI}	var $\widehat{\Delta y^\star}$	w_{GDP}	w_{GDI}	var $\widehat{\Delta y^\star}$
Final Curr. Qtrly.	**0.32**	**0.68**	**24.77**	**0.68**	**0.32**	**25.44**	**0.30**	**0.65**	**3.49**	**0.70**	**0.35**	**4.36**
	(0.25)	(0.25)		(0.25)	(0.25)		(0.23)	(0.25)		(0.23)	(0.25)	
First Annual	**0.49**	**0.50**	**28.06**	**0.51**	**0.50**	**29.24**	**0.46**	**0.47**	**3.53**	**0.54**	**0.53**	**4.70**
	(0.20)	(0.20)		(0.20)	(0.20)		(0.19)	(0.19)		(0.19)	(0.19)	
Second Annual	**0.53**	**0.45**	**27.28**	**0.47**	**0.55**	**28.56**	**0.49**	**0.42**	**3.20**	**0.51**	**0.58**	**4.75**
	(0.17)	(0.17)		(0.17)	(0.17)		(0.16)	(0.15)		(0.16)	(0.15)	
Third Annual	**0.56**	**0.42**	**27.29**	**0.44**	**0.58**	**29.02**	**0.49**	**0.37**	**2.98**	**0.51**	**0.63**	**5.44**
	(0.14)	(0.14)		(0.14)	(0.14)		(0.13)	(0.12)		(0.13)	(0.12)	
Latest Available	**0.63**	**0.34**	**30.77**	**0.37**	**0.66**	**33.20**	**0.50**	**0.27**	**2.39**	**0.50**	**0.73**	**6.41**
	(0.13)	(0.12)		(0.13)	(0.12)		(0.12)	(0.09)		(0.12)	(0.09)	

Notes

[1]National Income accountants face two fundamental problems. First, they must define an interesting and useful measure of aggregate economic activity, and second, they must design methods for estimating the value of that measure, taking the definition as fixed. Our concern in this paper is with the second issue, using the definition of economic activity traditionally employed by National Income accountants. It is a value-added measure, and the private sector component is restricted to marketed economic activity for the most part, excluding non-market activities such home production and changes in natural resources. For more discussion and references, see Stone's Nobel Memorial lecture, Sir Richard Stone (1984).

[2]Our terminology follows N. Gregory Mankiw and Matthew D. Shapiro (1986); see also Mankiw, David E. Runkle and Shapiro (1984).

[3]See Martin R. Weale (1992), and the related work in E. Phillip Howrey (2003), Weale (1985), and Richard J. Smith, Weale, and Steven E. Satchell (1998).

[4]Since BEA also produces vintages of estimates it is possible to examine the change in a particular period's estimate across vintages. We do some of this in this paper; much more of this analysis is conducted in BEA's revision studies; see for example, Dennis J. Fixler and Bruce T. Grimm (2002).

[5]All data have been deflated by the GDP deflator. The model includes the 1984 break in the variance of "true" unobserved GDP growth.

[6]We should note that our efficiency assumption is weaker than some others that have been tested in the literature, such as those in Karen E. Dynan and Douglas W. Elmendorf (2001) and Dennis J. Fixler and Bruce T. Grimm (2003). We only assume

that the estimates are efficient with respect to the internal information used to compute them, not with respect to the entire universe of available information - we do not consider efficiency with respect to the slope of the yield curve, stock prices, and so on.

[7]A reader may question whether it is possible to compute an efficient estimate of $\Delta y_t^\star$ given that it is unobserved. A couple of things should be kept in mind. First, though $\Delta y_t^\star$ itself is unobserved, it is defined quite precisely - see footnote 1. Second, BEA and statisticians in general draw on a large stock of knowledge about the data they employ, and it's reliability. More reliable data sources are generally given greater weight, and less reliable data sources less weight; through such procedures it may be possible to produce estimates that are close to efficient even though $\Delta y_t^\star$ is never observed. To illustrate, suppose that the source data used to compute a component of GDP is contaminated with sampling error, and the variance of the sampling error is known (as is often the case); then standard procedures may be employed to downweight the estimate in proportion to the variance of the sampling error, producing an efficient estimate for that component even though it's true value is never observed.

[8]In an extended set of additional results we offer as an Appendix, we work through and estimate models that allow for serial correlation of arbitrary linear form in GDP and GDI. The main points of the paper carry through in this setting, and the empirical estimates with dynamics are similar to the empirical estimates of the static models presented here.

[9]Additional covariance between the estimates may arise from correlation between the measurement errors ε_t^1 and ε_t^2. We have worked through this case, and found the formulas to be slightly less transparent but similar to those reported; the main points about news vs. noise carry through in the setting. For this reason, and because our model is already underidentified without the inclusion of an additional parameter, we

have chosen to focus on the case of uncorrelated measurement errors.

[10]This equation remains correct when we relax the assumption $E\left(\Delta y_t^\star | \mathcal{F}_t^k\right) = \Delta y_t^\star$. The only difference lies in the interpretation of the parameters. With this assumption, σ^2 identifies the variance of "true" GDP growth. Without this assumption, σ^2 merely identifies $\mathrm{var}\left(E\left(\Delta y_t^\star | \mathcal{F}_t^1\right)\right) = \mathrm{var}\left(E\left(\Delta y_t^\star | \mathcal{F}_t^2\right)\right)$, which must be less than the variance of "true" GDP growth.

[11]The definition of national income changed in BEA's 2003 benchmark revision, but this only served to reshuffle some items within GDI; continuity with earlier GDI vintages was maintained. For more information on how GDP, GDI and their components are constructed over our sample, see the October 2002 Survey of Current Business and references there-in, or visit www.bea.gov.

[12]We choose to focus on nominal data in our combining exercises because BEA does not produce a deflator for GDI; after the combined estimates have been computed they can be deflated by any deflator the researcher deems appropriate, for example in Figure 1 we chose the GDP deflator. We did experiment with deflating both GDP and GDI by the GDP deflator and then combining them; this process gave similar weighting results to those reported.

[13]Evidence supporting this choice for the break point is provided below.

[14]For the pure noise model, σ^2 identifies the variance of "true" GDP growth if $E\left(\Delta y_t^\star | \mathcal{F}_t^k\right) = \Delta y_t^\star$; without this assumption σ^2 is only a lower bound on the variance of "true" GDP growth, similar to the pure news case.

[15]We thank Mark W. Watson for making us aware of the Minimax approach, and providing some examples which essentially showed that the weightings will end up somewhere between the weights dictated by the pure news and pure noise models, with the

strength of priors dictating where the weights fall. See Mark W. Watson (1987) and E. L. Lehmann and George Casella (1998) for an example and description of the Minimax approach.

[16]These weights can be derived through application of L'Hopital's rule.

[17]Weale (1992) allowed for covariance between the measurement errors ε_t^1 and ε_t^2. This has no impact on the weights when they are constrained to sum to one.

Figure 1:
1999-2002 Growth Rates of Real GDP, Real GDI, and
Estimated "True" Growth Rate (under news model)

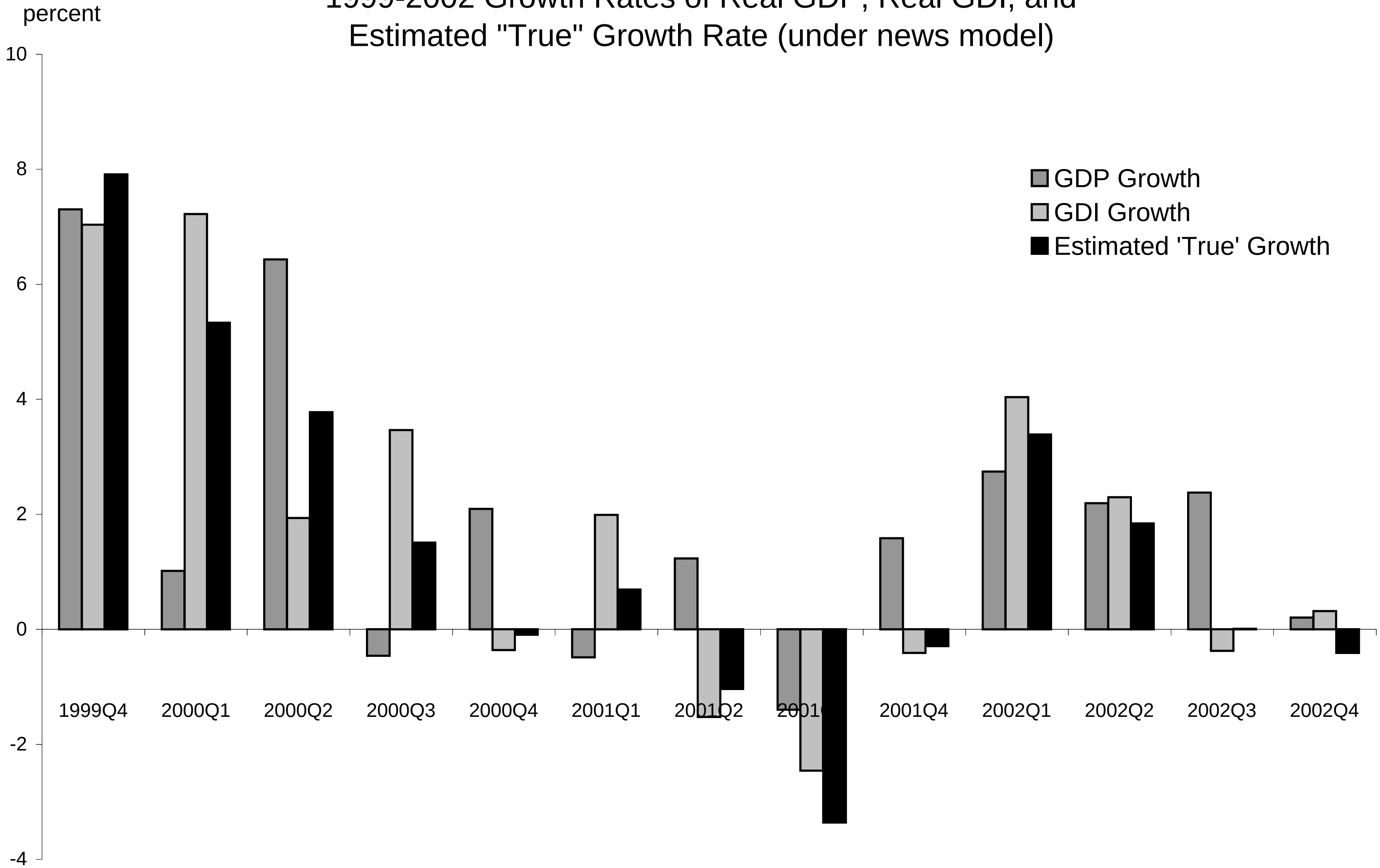

Figure 2:
1978 to 2002 Growth Rates of Nominal GDP and GDI,
Latest Available data as of August 2005