

**Aggregation Issues in Integrating and Accelerating BEA's Accounts:  
Improved Methods for Calculating GDP by Industry**

*Brian Moyer, Marshall Reinsdorf and Robert Yuskavage*

*Bureau of Economic Analysis*  
1441 L St. NW  
Washington DC 20230

February 2004

The *GDP by Industry Accounts* prepared by the Bureau of Economic Analysis (BEA) are frequently used to study structural change and sources of growth in the U.S. economy, to compare U.S. industrial performance with that of other countries, and to assess the contributions of industries and sectors to aggregate productivity growth. By providing annual estimates of nominal and real gross output, intermediate inputs, and value added for 66 industries, these accounts allow researchers to understand changes over time in the relative importance of industries. The nominal (current-dollar) value added estimates provide measures of industry size relative to GDP, and the real value added estimates provide measures of industry contributions to real GDP growth.

Aggregate measures of real GDP growth obtained from the GDP by Industry Accounts, however, often differ from the featured measure of real GDP growth obtained from the National Income and Product Accounts (NIPAs). Because these differences have raised concerns among researchers about the consistency of the industry and national economic accounts, BEA is working on a more complete integration of these accounts that would reduce or eliminate existing discrepancies. BEA is also investigating whether changes in methodology can reduce discrepancies between the sum of the industry contributions and real GDP growth from the NIPAs. One of our most important findings is that the same “exact contributions” formula used to calculate the contributions of final expenditures to real GDP growth in the NIPAs can also be used to calculate industry contributions based on value added.

In this paper, we describe some of the causes of discrepancies between estimates based on the GDP by Industry Accounts and estimates based on the NIPAs, and we identify several options for bringing BEA’s aggregate real output measures into closer alignment. We investigate reasons for the differences between the growth of real GDP and the sum of the industry contributions to real growth, including the treatment of the statistical discrepancy, differences in the data sources and methods used for the expenditures and industry (production) approaches to measuring GDP, deflation and aggregation methods, and the contributions formula itself. Reasons for the nominal statistical discrepancy are beyond the scope of this paper.

This paper also tests the feasibility of short-run and long-run options for bringing the aggregate real output measures into closer alignment using newly developed datasets. This research is one of the goals in BEA’s multi-year strategic plan for better integrating the industry and national accounts. Possible options, which are described in the last section of the paper, include partial or full integration of the different approaches to measuring GDP, modifications to

the contributions formula, and changes in presentation of the estimates. This paper also identifies improvements in source data that are needed to achieve more highly integrated national and industry economic accounts.

An important conclusion from this paper is that differences in source data, combined with differences in methodology, account for most of the difference in aggregate real output growth rates; very little of the difference is attributable to the treatment of the statistical discrepancy, differences in aggregation methods, or the contributions formula. In fact, this paper demonstrates that with consistent data, the Fisher-Ideal aggregation procedure used by BEA to measure real GDP yields the same estimate when real GDP is obtained by aggregating value added across industries as when real GDP is measured by aggregating final uses of commodities. Thus, two major approaches to measuring real GDP—the “expenditures” approach used in the NIPAs and the “production” or “industry” approach used in the Industry Accounts—give the same answer under certain conditions. This result also leads to the finding that the NIPA “exact contributions” formula can also be used for GDP by Industry. Although these results imply that some sources of discrepancy could be eliminated, accomplishing this would require improvements in industry source data to go along with the more integrated estimation framework.

The remainder of the paper is presented in four sections. Section I provides background on the GDP by Industry Accounts and the magnitude of the existing discrepancies. Section II describes the sources of the existing discrepancies for nominal shares and real contributions. These factors include methodology and source data, deflation and aggregation procedures, the treatment of the statistical discrepancy, and the current contributions formula. Section III presents the empirical results, including tests of the relative importance of the factors described above. This section also describes how the research datasets were developed and the ways they were used to evaluate the various sources of difference. Section IV is a summary and conclusion that describes possible solutions to the discrepancies, options for implementation, and directions for future research on integration of the industry and national accounts.

## **I. Discrepancies between the Industry and National Accounts**

The industry estimates of nominal value added from the GDP by Industry Accounts are largely derived from the income-side industry estimates in the NIPAs. The total for gross domestic income (GDI) in the NIPAs, however, differs from the featured expenditure-based

estimate of gross domestic product (GDP) by an amount known as the statistical discrepancy. Therefore, to balance GDP by Industry summed over all industries with the expenditure-based estimate of gross domestic product (GDP) in the NIPAs, the industry estimates include the statistical discrepancy as a separate “industry”.

As a result of the statistical discrepancy, industry shares of nominal GDP rarely sum to unity, and in recent years the statistical discrepancy has occasionally exceeded one percent of GDP in absolute value. Furthermore, for several reasons real output for all industries combined from the GDP by Industry Accounts usually differs from the product-side estimate of real GDP; indeed, in some years the growth rates differ by several tenths of a percentage point. This is a major reason why the published industry contributions to real GDP growth do not necessarily sum to the growth in real GDP. These discrepancies cause problems for researchers who are using the real value added by industry estimates for studying industry performance and contributions to productivity growth. (For a recent example, see Faruqi et. al.)

To illustrate the magnitude of the problem, table 1 presents the published shares of nominal GDP and contributions to real GDP growth for industry groups and higher-level aggregates for 1999-2001.<sup>1</sup> The industry groups shown are aggregates of the more detailed, generally two-digit Standard Industrial Classification (SIC) industries found in the regularly-published GDP by Industry Accounts. The higher-level aggregates include private industries, private goods-producing industries, private services-producing industries, and government. Table 1 also presents shares and contributions that are “Not allocated by industry”, which consist of the statistical discrepancy and “other” amounts not allocated by industry.<sup>2</sup> Since the statistical discrepancy was negative in each year, industry group contributions sum to more than 100 percent of GDP.

For shares of nominal GDP, the amount “not allocated by industry” consists only of the statistical discrepancy. For contributions to real GDP growth, however, the amount “not allocated by industry” represents the combined effects of the real statistical discrepancy and other factors, such as differences in source data, methodology, aggregation procedures, and the contributions formula itself. These other factors account for some of the difference between real

---

<sup>1</sup> Revised estimates of GDP by Industry that are consistent with the 2003 NIPA comprehensive revision and that are classified on the 1997 North American Classification System (NAICS) basis will be available in June 2004.

<sup>2</sup> For a more detailed description of the amounts not allocated by industry, see the box entitled “Nonadditivity of Chained Dollars and “Not Allocated by Industry” in the GDP-by-Industry Accounts”, in McCahill and Moyer.

GDP growth and the sum of the industry contributions. The statistical discrepancy made an unusually large contribution to real GDP growth in 2000 (-0.94 percentage points) because of the large increase in the nominal statistical discrepancy between 1999 and 2000. In 1999, other factors contributed -0.27 percentage points, primarily reflecting faster growth in real GDP by Industry for “all industries” than in the published real GDP growth from the NIPAs.

**Table 1**  
**Industry Group Shares of GDP and Contributions to Real GDP Growth, 1999-2001**

Industry Group	Shares			Contributions		
	1999	2000	2001	1999	2000	2001
Gross domestic product	100.0	100.0	100.0	4.1	3.8	0.3
Private industries <sup>a</sup>	87.6	87.6	87.3	4.21	3.42	0.34
Private goods-producing industries	23.1	22.9	21.6	1.06	0.83	-0.96
Agriculture, forestry, and fishing	1.4	1.4	1.4	0.09	0.11	-0.02
Mining	1.1	1.4	1.1	-0.05	-0.13	0.06
Construction	4.6	4.7	4.8	0.23	0.13	-0.08
Manufacturing	16.0	15.5	14.1	0.78	0.75	-0.93
Durable goods	9.2	9.0	8.1	0.60	0.92	-0.47
Nondurable goods	6.8	6.5	6.1	0.19	-0.15	-0.46
Private services-producing industries	64.9	66.0	66.8	3.23	3.54	1.15
Transportation and public utilities	8.3	8.2	8.1	0.60	0.56	-0.01
Transportation	3.3	3.2	3.0	0.14	0.17	-0.14
Communications	2.8	2.8	2.9	0.28	0.34	0.35
Electric, gas, and sanitary services	2.3	2.2	2.2	0.18	0.05	-0.20
Wholesale trade	7.0	7.1	6.8	0.47	0.41	-0.01
Retail trade	9.0	9.0	9.2	0.52	0.67	0.42
Finance, insurance, and real estate	19.4	20.1	20.6	0.79	1.21	0.56
Services	21.3	21.5	22.1	0.85	0.69	0.20
Government	12.4	12.4	12.7	0.16	0.33	0.21
Not allocated by industry	-0.4	-1.3	-1.2	-0.35	-1.00	-0.18
Statistical discrepancy	-0.4	-1.3	-1.2	-0.08	-0.94	0.14
Other	NA	NA	NA	-0.27	-0.05	-0.32

a. Includes the statistical discrepancy

## II. Sources of Discrepancies

This section describes the factors that contribute to the existing discrepancies for shares of nominal GDP and for contributions to real GDP growth. These sources of discrepancies include methodology and source data, deflation and aggregation procedures, the treatment of the statistical discrepancy, and the contributions formula currently used by BEA. Each of these sources of difference is described separately.

### A. Methodology and source data

Different methodologies can lead to different estimates of aggregate output levels and growth rates, as well as different estimates of the shares and contributions to growth of the components of aggregate output. BEA currently uses two approaches, the expenditures approach and the income approach, to measure GDP. The expenditures approach measures GDP as the sum of final uses of goods and services, which consist of personal consumption expenditures, gross private domestic investment, net exports of goods and services, and government consumption expenditures and gross investment. This approach provides a good framework for measuring real GDP because it relies on detailed current-dollar data that can be deflated by price indexes to compute quantity indexes. The income approach measures GDP as the sum of the costs incurred and incomes earned in production, including compensation of employees, gross operating surplus such as corporate profits, proprietors' income, capital consumption allowances, and net interest, and other charges against GDP such as taxes on production and imports. Real GDP is not estimated using the income approach, because the components of gross domestic income (GDI) cannot be separated into price and quantity components.

In addition to the expenditures and income approaches, the 1993 *System of National Accounts* (SNA93) identifies the *production approach* (also known as the *industry approach*) as a third way to measure GDP. In the production approach, GDP is calculated as the sum over all industries—including government—of gross output (sales) less intermediate inputs (purchases). With this method, real GDP can be computed using the double-deflation method as the difference between real gross output and real intermediate inputs for all industries. Although BEA does not use this approach to measure GDP, a variant of it is used for preparing the estimates of real value added by industry in the GDP by Industry Accounts.<sup>3</sup>

---

<sup>3</sup> The input-output (I-O) accounts compute nominal value added by industry using the production approach, but the total over all industries in the I-O accounts is benchmarked to the final expenditures estimate of GDP.

Figure 1 is a diagram of a highly-aggregated input-output (I-O) “use table” that can illustrate the three different approaches to measuring GDP.<sup>4</sup> Industries, final uses, and total commodity output are the major column descriptions, and commodities, value added, and total industry output are the major row descriptions. Because total commodity output equals total industry output, and because the same value of total intermediate uses is subtracted from both measures of gross output, final uses summed over all commodities equals value added summed over all industries. The expenditures approach to measuring GDP is the equivalent of summing final uses over each of the sub-categories (e.g., PCE) and each of the commodities (e.g., manufacturing). This is shown in the shaded column. The incomes approach to measuring GDP is the equivalent of summing each of the value-added components (such as “compensation”) over all industries. The production approach is equivalent to summing each industry’s total value-added over all industries. This is shown in the shaded row.

In concept, these three approaches yield the same measure of GDP, but in practice they generally differ because they use source data that are not entirely consistent. The source data for implementing the expenditures approach are derived largely from Census Bureau business surveys, but allocations of some commodities between final uses and intermediate uses are often based on the benchmark I-O accounts for economic census years. The source data for the incomes approach are largely derived from administrative records such as business tax returns. Census Bureau business surveys also provide source data that could be used to measure gross output in the production approach, but the allocations between intermediate uses and value added would be more reliant on the I-O accounts than are the estimates of final demand under the expenditures approach. While the production approach could be used to measure both nominal and real GDP, major improvements would be needed in the source data for gross output for selected industries, price indexes, and intermediate inputs, especially purchased services. BEA has not attempted to prepare independent measures of GDP using the production approach.

---

<sup>4</sup> For a description of BEA’s benchmark input-output accounts, see Lawson et. al.

**Figure 1. -- Input-Output Use Table**

		INDUSTRIES										FINAL USES (GDP)							TOTAL COMMODITY OUTPUT
		Agriculture	Mining	Construction	Manufacturing	Transportation	Trade	Finance	Services	Other	Total Intermediate Use	PCE	PFI	CBI	X	M	GOVT	GDP	
COMMODITIES	Agriculture																		
	Minerals																		
	Construction																		
	Manufacturing																		
	Transportation																		
	Trade																		
	Finance																		
	Services																		
	Other																		
	Noncomparable imports																		
	Total Intermediate inputs																		
VALUE ADDED	Compensation																		
	IBT																		
	Other value added																		
	Total																		
<b>TOTAL INDUSTRY OUTPUT</b>																			

Abbreviations: PCE–personal consumption expenditures; PFI–private fixed investment; CBI–change in business inventories; IBT–indirect business taxes.



## B. Deflation and aggregation procedures

### 1. Theoretical overview

The use table shown in figure 1 is part of an integrated estimation framework in the I-O accounts that yields both a production approach estimate of real GDP and an expenditures approach estimate of real GDP. The other components of this estimation framework are the make table and the deflators for the commodities shown in the make table and the use table. The make table shows the value of each primary or secondary commodity produced by each industry, while the use table shows the use of each commodity as an intermediate input by each industry. To estimate real GDP using either the production approach or the expenditures approach, the current-dollar values in the make and use tables--which are measured at producers' prices--must be deflated by indexes of producers' prices for each commodity.<sup>5</sup>

In the absence of data inconsistencies, the production approach estimate of nominal GDP calculated from the make and use tables agrees with the expenditures approach estimate because the two approaches differ only in the order in which they combine the elements of the “make” and “use” tables from the I-O accounts. The production approach first aggregates over commodities within each industry, and then aggregates over industries. Letting  $V_{cit}$  represent the production of commodity  $c$  by industry  $i$  in year  $t$  from the make table, the industry's gross output  $g_{it}$  equals:

$$g_{it} = \sum_c V_{cit}. \quad (1)$$

Letting  $U_{cit}$  represent the use of commodity  $c$  by industry  $i$  in year  $t$  from the use table, for industry  $i$  in year  $t$  the total use of intermediate inputs  $m_{it}$  equals:

$$m_{it} = \sum_c U_{cit}. \quad (2)$$

The production approach estimate of nominal GDP is, then:

$$\begin{aligned} TVA_t &= \sum_i (g_{it} - m_{it}) \\ &= \sum_i VA_{it} \end{aligned} \quad (3)$$

where  $VA_{it}$  represents value added of industry  $i$  in period  $t$

The expenditures approach first aggregates commodity gross output net of intermediate uses over industries to obtain the final use of each commodity in GDP, and then sums over all commodities. Final uses  $e_{ct}$  of commodity  $c$  are:

$$e_{ct} = \sum_i [V_{cit} - U_{cit}] \quad (4)$$

The expenditures approach estimate of nominal GDP is then:

$$GDP_t = \sum_c e_{ct}. \quad (5)$$

The production approach will also agree with the expenditures approach in estimating real GDP provided that the deflator for any commodity is the same wherever that commodity is used (this assumption is more likely to hold if commodities and their deflators are defined at a high level of detail). Real GDP growth is defined as the growth rate of a Fisher index calculated from Laspeyres and Paasche constant-dollar estimates of GDP growth.

The GDP by Industry Accounts calculate an industry's real value added as the difference between real gross output and real intermediate inputs, a method known as "double-deflation". Because Fisher indexes lack the property of *consistency in aggregation*, Fisher measures of value added must be computed from separate Laspeyres and Paasche measures of gross output and intermediate inputs, not from Fisher measures of output and inputs.<sup>6</sup> The Fisher index for real value added in an industry is therefore calculated as the geometric mean of one value added index based on Laspeyres double-deflation and another index based on Paasche double-deflation.

Real value added for "all industries" — the production approach estimate of real GDP — is an aggregate Fisher quantity index for 62 private industries and four types of government. Yuskavage (1996, p. 142) explains how the aggregate Fisher index is calculated. Separate Laspeyres and Paasche indexes are computed for the aggregate of all industries, resulting in two sets of estimates of economy-wide real gross output and economy-wide real intermediate inputs. Next, Laspeyres and Paasche indexes of aggregate value added are computed by subtracting

<sup>5</sup> Use tables but not make tables are also available valued at purchasers' prices.

<sup>6</sup> An index number formula is consistent in aggregation if calculating lower level aggregates using the formula and then combining these lower level aggregates into a top level aggregate using that same formula yields the same result as using the formula just once to calculate the top level aggregate directly from the detailed components (Vartia, 1976, p. 124.) The Fisher formula is not consistent in aggregation, though Diewert (1978) shows that it is approximately consistent in aggregation.

economy-wide intermediate inputs from economy-wide gross output, then averaged to obtain the aggregate Fisher index.

To calculate a Laspeyres constant-dollar estimate of GDP in time  $t$ , we first deflate each  $V_{cit}$  and each  $U_{cit}$  by  $r_{ct}$ , the deflator from time  $t-1$  to time  $t$  for commodity  $c$ . To obtain the production approach Laspeyres index, we then use these deflated values in equations (1) through (3), and to obtain the expenditures approach Laspeyres index we use these values in equations (4) and (5). The equivalence of the production and expenditures approaches then follows from the fact that they both combine the same elements of the deflated make and use table to compute the numerator of the Laspeyres index.

Similarly, to obtain the Paasche constant-dollar estimates of GDP for time  $t$ , we reflate each  $V_{ci,t-1}$  and each  $U_{ci,t-1}$  by  $r_{ct}$  and then apply equations (1) through (3) for the production approach or equations (4) and (5) for the expenditures approach. The order of addition of the elements of the make and use tables is again the only difference between the expenditures approach and the production approach; in particular, both approaches compute the denominator of the Paasche index as the same combination of the entries in the reflatd make and use tables. Whether the production approach or the expenditures approach is used therefore has no effect on the Laspeyres and Paasche indexes on which the Fisher index depends.

Since — given the assumptions of consistent data and uniform deflators — real GDP growth is the same measured by the production approach as it is measured by the expenditures approach, use of double deflation does not itself cause a discrepancy between the measure of real GDP from the Industry accounts and the measure of real GDP in the NIPAs. In theory, a decomposition of real GDP into industry contributions that add up exactly to the NIPA measure requires only a way to identify the contribution of each industry to a Fisher index aggregate of industries.

Using expenditures on final uses for weighting purposes, the Laspeyres price index for GDP is defined as:

$$L^P = \frac{\sum_c e_{ct-1} r_{ct}}{\sum_c e_{ct-1}}. \quad (6)$$

Similarly, the Paasche price index is:

$$P^P = \frac{\sum_c e_{ct}}{\sum_c e_{ct} / r_{ct}}. \quad (7)$$

The Fisher price index  $F^P$  is defined as the geometric mean of  $L^P$  and  $P^P$ . Finally, the Fisher quantity index may be defined as the expenditures change deflated by the Fisher price index. Hence, the change in real GDP at time  $t$  equals the change in nominal GDP deflated by  $F^P$ :

$$F^Q = \frac{GDP_t}{GDP_{t-1}} \frac{1}{F^P}. \quad (8)$$

The following proposition shows how to express  $F^Q$  as the change in the sum over commodities of final uses, and also as the change in the sum over industries of value added. The method requires both deflated make and use tables from period  $t$  and reflatd make and use tables from period  $t-1$ . Each deflated or reflatd make or use table effectively holds prices constant at an average of their level in period  $t-1$  and their deflated level in period  $t$ , where  $F^P$  is taken to be the appropriate deflator. Exactly additive commodity contributions to the change in real GDP are implied by the final uses of commodities measured at these constant prices, and exactly additive industry contributions are implied by the constant-price measures of value added.

**PROPOSITION 1:** Define  $h_{ct}$  as the harmonic mean of  $r_{ct}$  and  $F^P$ , the Fisher price index for the expenditure-approach estimate of GDP:

$$h_{ct} \equiv \frac{2}{1/r_{ct} + 1/F^P}. \quad (9)$$

Also, define  $a_{ct}$  as the arithmetic mean of  $r_{ct}/F^P$  and 1:

$$a_{ct} \equiv \frac{r_{ct}/F^P + 1}{2}. \quad (10)$$

Then:

(a) The Fisher estimate of real GDP equals

$$F^Q = \frac{\sum_c e_{ct} / h_{ct}}{\sum_c e_{ct-1} a_{ct}}. \quad (11)$$

(b) The additive contribution  $C_\gamma$  of the arbitrary commodity  $\gamma$  to the change in  $F^Q$  is:

$$C_\gamma = \frac{e_{\gamma t}/h_{\gamma t} - e_{\gamma t-1} a_{\gamma t}}{\sum_c e_{ct-1} a_{ct}}. \quad (12)$$

(c) The additive contribution  $\hat{C}_j$  of the arbitrary industry  $j$  to the change in  $F^Q$  is:

$$\hat{C}_j = \frac{\sum_c V_{cjt}/h_{ct} - \sum_c U_{cjt}/h_{ct} - [\sum_c V_{cjt-1} a_{ct} - \sum_c U_{cjt-1} a_{ct}]}{\sum_i [\sum_c V_{cjt-1} a_{ct} - \sum_c U_{cjt-1} a_{ct}]}. \quad (13)$$

**PROOF:** To prove part (a), note that by equation (6),

$$\sum_c e_{ct-1} a_{ct} = GDP_{t-1} (L^P/F^P + 1)/2. \quad (14)$$

From equation (7),

$$\sum_c e_{ct}/h_{ct} = GDP_t (1/P^P + 1/F^P)/2. \quad (15)$$

Therefore,

$$\begin{aligned} \frac{\sum_c e_{ct}/h_{ct}}{\sum_c e_{ct-1} a_{ct}} &= \frac{GDP_t}{GDP_{t-1}} \frac{1/P^P + 1/F^P}{L^P/F^P + 1} \\ &= \frac{GDP_t}{GDP_{t-1}} \frac{P^P + F^P}{P^P F^P (L^P/F^P + 1)} \\ &= \frac{GDP_t}{GDP_{t-1}} \frac{P^P + F^P}{F^P (P^P + F^P)} \\ &= F^Q. \end{aligned} \quad (16)$$

Part (b) of proposition 1 is an immediate corollary of part (a). Substituting from equation (4) for  $e_{ct}$  in part (a) of proposition 1 then rearranging yields the equation in part (c).

Using Proposition 1, we deflate all the entries for each commodity  $c$  in the use and make tables for period  $t$  by  $h_{ct}$  and we reflate all the entries for commodity  $c$  in the use and make tables for period  $t-1$  by  $a_{ct}$ . Summed over industries, these adjusted use and make tables yield the *commodity* contributions to change of Proposition 1:

$$C_{\gamma} = \frac{[\sum_i V_{\gamma it} - \sum_i U_{\gamma it}] / h_{\gamma t} - [\sum_i V_{\gamma it-1} - \sum_i U_{\gamma it-1}] a_{\gamma t}}{\sum_c [\sum_i V_{cit-1} - \sum_i U_{cit-1}] a_{ct}}. \quad (17)$$

When the adjusted entries in the make and use tables are instead summed over commodities to obtain adjusted values of  $VA_{it}$ , they provide exact *industry* contributions to the change in a production approach estimate of real GDP,  $\hat{C}_i$ .

Note that the formula for contributions to change has the price index for the aggregate to be decomposed,  $F^P$ , as one of its arguments. This dependence on the price index of the aggregate to be decomposed means that the relative sizes of contributions can change if the definition of the aggregate is altered. The contributions to change in real GDP depend on  $F^P$  because they value quantity changes for commodities based on a price vector that is a weighted average of time  $t$  prices and time  $t-1$  prices, where the weight given the time  $t$  prices is inversely proportional to  $F^P$ .

## 2. Differences between Theory and Practice

The agreement that exists in theory between the expenditures approach estimate of real GDP and the production approach estimate is difficult to achieve in practice because of inconsistencies in source data and in deflators constructed from different kinds of prices. Even within the fully-integrated framework of the input-output accounts, estimates must be balanced in constant prices as well as in current prices. This balancing process often raises thorny practical issues because of the need to reconcile underlying inconsistencies in both nominal values and price indexes. Agreement between the currently-used expenditures approach estimate of real GDP from the NIPAs and the production approach estimate from the GDP by Industry Accounts is likewise very difficult to achieve because the source data used for the two approaches are not completely consistent.<sup>7</sup>

Nevertheless, differences in the quality and detail of available source data most likely render the NIPA expenditures approach more accurate for measuring GDP than an integrated

---

<sup>7</sup> Similar data inconsistencies cause problems for those countries that try to combine both approaches.

expenditures/production approach might be. In particular, for many commodities the NIPAs can use data that directly measure narrow categories of final expenditures, eliminating the need to rely on input-output relationships for deriving final uses from total commodity supply.<sup>8</sup> Also, in the NIPAs the components of final expenditures and the price indexes used to compute real GDP are generally quite detailed, but in the Industry Accounts consistent and detailed data on commodity output and prices are available just for the manufacturing industries; for other kinds of industries, output data are often not detailed or not completely consistent. For intermediate inputs, detail is quite extensive and consistent across industries, but these data are not as timely as the data on the components of gross output.

The use of less detailed and less timely data in parts of the GDP by Industry Accounts is not the only source of difference in real estimates. Price indexes also differ because the GDP by Industry Accounts use producers' price, while the NIPAs use purchasers' prices, which include wholesale and retail trade margins and transport costs. Price indexes used for deflation in the NIPAs, such as components of the BLS consumer price index (CPI) and the BLS export and import price indexes, generally reflect purchaser price concepts and thus can be used directly.<sup>9</sup> These differences in deflation procedures mean that  $F^P$  in the GDP by Industry Accounts — which plays a critical role in the contributions formulas of Proposition 1 — can be expected to deviate slightly from the price index for GDP in the NIPAs.

Another kind of discrepancy in the published GDP by Industry Accounts is that the published industry contributions to change in real GDP generally do not sum to even the (unpublished) production approach estimate of real GDP growth. Calculating industry contributions to the production approach estimate of real GDP is a difficult problem because Fisher indexes are not consistent in aggregation. This means that the total over all industries of the Fisher index estimate of real value added in each industry is algebraically different from the production approach estimate of real GDP. Hence, an industry's contribution cannot be calculated simply by dividing its real value added by the production approach estimate of real GDP. The difference between the sum of the published industry contributions and the actual change in NIPA real GDP is known as the amount "Not allocated by industry" (NAI). Data

---

<sup>8</sup> For some commodities, however, such as restaurant meals and beverages and air passenger transportation, an assumption must be made that relationships between total supply and final uses have not changed since the latest benchmark I-O accounts.

<sup>9</sup> However, producer price indexes are used for some items, such as some business investment in equipment.

inconsistencies—including the statistical discrepancy—contribute to the NAI residual, but with the formula currently used to compute contributions to change, this residual would exist even in the absence of data inconsistencies.

A formula for contributions to change that would eliminate the NAI residual in the absence of data inconsistencies, however, is given by equation (13). This formula extends the approach that the NIPAs use for contributions to change in a Fisher index to a new application, double-deflation, an idea that was suggested by Dumagan (2002). (For additional background on the NIPA formula for contributions to change in real GDP, see Reinsdorf, Diewert and Ehemann (2002).) To use equation (13) in practice, however, requires some algebraic manipulation because the GDP by Industry Accounts currently do not include complete make and use tables. (Make and use tables are scheduled to become available in June 2004 in data sets that “partially integrate” the GDP by Industry Accounts and the I-O Accounts.) The appendix shows how to express  $\hat{C}_i$  as a function of data that are available in the GDP by Industry Accounts, in particular, Laspeyres and Paasche indexes for industry gross output and intermediate inputs.

### *C. Statistical discrepancy*

The statistical discrepancy is defined as current-dollar GDP less GDI. It is recorded in the NIPA's as an “income” component that reconciles the income side with the product side of the accounts. It arises because the two sides are estimated using independent and imperfect data. For the GDP by industry estimates, which are derived from the income side of the accounts, the statistical discrepancy is treated as an industry, such that nominal GDP by industry sums to nominal GDP. This balancing role for the statistical discrepancy in GDP by industry carries over directly from its balancing role in the NIPA's. The real statistical discrepancy is computed by deflating the nominal (current-dollar) statistical discrepancy with the implicit price deflator (IPD) for the business sector in GDP. This choice for a deflator reflects BEA's view that the source data inconsistencies underlying the statistical discrepancy are most likely located in a broad spectrum of private business sector industries. Otherwise, assumptions would need to be made about which industries are most likely affected by this discrepancy.

One of the most important uses of the nominal GDP by industry estimates is to calculate an industry's share of nominal GDP. These shares can be used to determine the relative size of an industry at a point in time, and how relative sizes are changing over time. A non-zero statistical discrepancy clouds the interpretation of these shares because some portion of GDP is



not accounted for in the value added of a specific industry. The statistical discrepancy indicates that the nominal value added for at least one industry is either too high or too low, relative to the final expenditures estimate of GDP. This problem is compounded when the statistical discrepancy is large and volatile, as it has been for recent years. Estimates of industry contributions to real GDP growth are also affected to the extent that the estimates of nominal value added growth are in error. In addition, because the statistical discrepancy is treated as industry, it is included in the calculation of real value added for “all industries.”

#### *D. Contributions formula*

The formula that is currently used for the published industry contributions to real GDP change is a Laspeyres approximation. This formula computes an industry’s contribution to the growth in an aggregate as the industry’s weighted growth rate, with the weight equal to the industry’s share of aggregate nominal value added in the first period. Aside from its computational simplicity, this formula avoids complications associated with including the statistical discrepancy as an industry. This discrepancy can change sign from one year to the next, making the use of the exact contributions formula very difficult. While the current contributions formula provides a close approximation to the exact contributions, it does not capture changes in shares between periods, and is not consistent with the procedure used to compute the Fisher quantity indexes for value added. Section B demonstrates, however, that the same exact contributions formula used for the NIPA’s can be used for GDP by Industry if the statistical discrepancy is not present and if source data inconsistencies are minimized, resulting in close agreement in aggregate growth rates.

### **III. Empirical Results**

This section presents the empirical results, including tests of the relative importance of the factors described above. This section also describes the datasets that were developed for this research and how these datasets were used to evaluate the various sources of difference. The empirical work was designed to assess the relative importance of several of the sources of difference described above. These results are presented in three sub-sections: Methodology and aggregation procedures, source data consistency (including the role of the statistical discrepancy), and the contributions formula.

### *A. Methodology and aggregation procedures*

One possible reason for the observed differences in aggregate growth rates and contributions is the use of different estimation methodologies and aggregation procedures. Both the published GDP by Industry and the NIPA's use Fisher aggregation procedures, but the estimation frameworks are quite different. Even if source data inconsistencies could be entirely eliminated, and if the same contributions formulas were used, it is not clear that aggregation over the existing GDP by Industry variant of the production approach would yield the same results as the NIPA final expenditures approach. A previous section has demonstrated, however, that consistent source data used in a consistent framework should yield the same aggregate indexes.

In order to test the impact of these possible sources of difference, an experimental "conceptually ideal" database was developed from the published annual input-output (I-O) accounts for 1998 and 1999. Nominal make and use tables were prepared at the summary level for 95 commodities and industries, and composite Fisher price indexes were computed for each commodity from detail underlying the GDP by Industry Accounts. As a result, the same price index was used to deflate a commodity regardless of whether it was consumed in final uses or intermediate uses. In addition, current-dollar source data were consistent among total supply, intermediate use and final use because of the use of balanced use and make tables. The 1999 tables were expressed in 1998 prices and the 1998 tables were expressed in 1999 prices in order to compute the necessary Laspeyres and Paasche quantity indexes for value added over industries and final uses over commodities.

In this experimental database, real growth rates are the same using both the expenditures and production approaches to measuring GDP (4.0 percent). Industry value added contributions based on the production approach sum exactly to real GDP growth using the exact Fisher formula. Table 2 shows the exact contributions to change calculated from the constant-price make and use tables for 1998 and 1999. Note that the price indexes used for these calculations are experimental, and may differ substantially from the price indexes used for the published estimates of real value added by industry and industry contributions to real GDP growth. Differences between the results in table 2 and the published estimates reflect other effects besides the use of the Fisher exact contributions formula.

By construction, the sum of the industry value added contributions to change in real GDP equal the sum of the commodity final use contributions. Table 2, however, demonstrates that

contributions can differ substantially between the commodity and the industry. Differences between commodity and industry contributions primarily reflect differential changes in the use of a commodity as an intermediate input and changes in an industry's use of intermediate inputs in its production process. For example, the construction commodity contributed much more to real GDP growth than the construction industry because an increased portion of the maintenance and repair construction commodity went to final uses in 1999, but little change took place in the construction industry's use of intermediate inputs. Also, the contribution to growth of manufacturing industries was below the contribution of manufactured commodities because the industries used relatively more intermediate inputs in 1999 but less of the production was used for intermediate purposes. On the other hand, mining commodities make a negative contribution to growth while the industry had a small positive contribution because of rising petroleum imports in 1999, which are a subtraction from final uses.

**Table 2.--Fisher Exact Contributions to Change in Real GDP by Commodity Final Use and by Industry Value Added, 1999**

	<b>Commodity Final Use</b>	<b>Industry Value Added</b>
Agriculture	0.0	0.1
Mining	-0.1	0.0
Construction	0.3	0.0
Manufacturing	0.6	0.4
Transportation, comm. & utilities	0.4	0.3
Trade	1.0	1.2
Finance, insurance & real estate	0.9	0.9
Services	1.0	1.1
Government	0.2	0.2
IVA	-0.3	-0.3
Noncomp. Imports & used goods	-0.1	NA
<b>TOTAL</b>	<b>4.0</b>	<b>4.0</b>

### *B. Source data consistency*

As described above, one possible reason for differences in real growth rates between GDP from the NIPA's and "all industries" from the industry accounts is the use of data from different sources within the industry accounts, along with the presence of the statistical discrepancy. For most industries, gross output is based on annual survey data collected by the Bureau of the Census, compensation of employees is based largely on data collected by the Bureau of Labor Statistics, and gross operating surplus is based largely on data reported on business income tax returns filed with the Internal Revenue Service. Because the same data reporting unit can be classified in different industries by different statistical agencies, inconsistencies often arise in the tabulated data, even at the two-digit SIC level. In addition, data are reported on corporate tax returns on a consolidated company basis rather than on an establishment basis. BEA converts the company-based estimates of corporate profits, corporate net interest, and corporate capital consumption allowances to establishment-based estimates using data on the employment of corporations. (See Yuskavage (2000) for a more detailed description and a discussion of the impact of these issues.)

For this research, BEA developed experimental industry time series of nominal and real gross output, intermediate inputs, and value added for 1992-2001 for 65 industries.<sup>10</sup> These estimates were consistent with the levels of both value added and gross output by industry from the 1992 benchmark I-O accounts, and did not include a statistical discrepancy. (This database was also used in research to test the feasibility of "partial integration" of BEA's industry accounts.) After first adjusting the levels in the 1992 benchmark I-O accounts to incorporate the definitions and conventions from the NIPA's and the GDP by Industry Accounts, nominal value added estimates were extrapolated annually using the published components of GDP by Industry for compensation of employees, gross operating surplus, and taxes on production and imports. The sum of these estimates over all industries was constrained to match nominal GDP from the NIPA's in each year, which eliminated the statistical discrepancy. Nominal gross output estimates were also benchmarked to the 1992 I-O accounts, and nominal intermediate inputs were obtained as a residual. Value-added quantity indexes were obtained for each industry using a modified double-deflation procedure that utilized the existing published chain-type price indexes for gross output and for intermediate inputs.

Aggregate “integrated” real value-added quantity indexes were computed for industry groups and for “all industries” using Fisher aggregation. Annual growth rates for “all industries” for the period 1993-2001 were compared with real growth rates for GDP and for “all industries” from the published GDP by Industry Accounts. The results are shown in table 3a. Relative to GDP, the mean error for the “integrated” estimates for 1993-01 is smaller than that for the “published” estimates (0.03 percentage points vs. 0.08 percentage points). The mean absolute error is about the same (0.19 points vs. 0.18 points). These results suggest that reducing the source data inconsistencies within the industry accounts would slightly reduce the differences in real growth rates between NIPA GDP and “all industries.” It is important to note, however, that the adjustments to improve consistency that were made for this research database are not as extensive as those that would be made in a formal “partial integration” methodology. As a result, these findings may understate the gain from using more consistent source data.

**Table 3a. – Annual Percent Changes in Aggregate Real Output Measures**

<u>Year</u>	NIPA <u>GDP</u> (1)	All Industries	
		<u>Published</u> <u>GDP by Ind.</u> (2)	<u>Integrated</u> <u>GDP by Ind.</u> (3)
1993	2.65	2.35	2.36
1994	4.04	3.90	3.87
1995	2.67	2.67	2.53
1996	3.57	3.82	3.84
1997	4.43	4.78	4.70
1998	4.28	4.28	4.17
1999	4.11	4.37	4.33
2000	3.75	3.75	3.72
2001	0.25	0.55	0.50
<u>Averages</u>			
1992-01	3.30	3.38	3.33
1992-00	3.69	3.74	3.69
1995-00	4.03	4.20	4.15

---

<sup>10</sup> These estimates were prepared by Abigail Kish of BEA’s Industry Economics Division. The revised version of the paper will include results from BEA’s integrated industry accounts scheduled for release in June. These revised estimates will incorporate the December 2003 comprehensive revision of the NIPA’s.

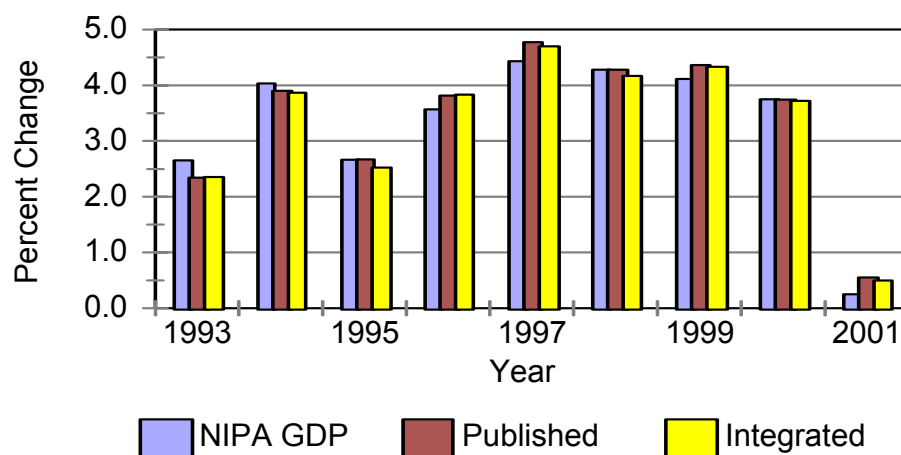
**Table 3b. -- Annual Percent Changes: All Industries less NIPA**

<u>Year</u>	Pub. All Ind.	Integ. All Ind.
	<u>Less NIPA</u> (1)	<u>less NIPA</u> (2)
1993	-0.31	-0.30
1994	-0.14	-0.17
1995	0.01	-0.14
1996	0.25	0.27
1997	0.35	0.27
1998	-0.00	-0.11
1999	0.25	0.22
2000	-0.01	-0.03
2001	0.30	0.25
Mean Error	0.078	0.029
MAE	0.179	0.194
<u>Averages</u>		
1992-01	0.078	0.029
1992-00	0.049	0.001
1995-00	0.168	0.123

Chart 1 below presents the annual percent change for NIPA GDP, published GDP by Industry for all industries, and integrated GDP by Industry for all industries. It is clear that most of the improvement resulting from use of the integrated estimates took place after 1996.

**Chart 1.-- NIPA vs. All Industries**

Published vs. Integrated GDP by Ind.



### C. Contributions formula

Table 4 presents the differences in industry contributions to real GDP growth for 1999-2001 caused by using a Laspeyres approximation rather than the Fisher exact contributions formula.<sup>11</sup> The format of this table is the same as table 1, which presented the published contributions of industry groups to real GDP growth. The amount “Not Allocated by Industry” consists of the statistical discrepancy and other factors, including the contributions formula. Exact contributions were calculated using the Laspeyres and Paasche quantity and price indexes underlying the published Fisher indexes.

**Table 4.- - Differences in Industry Group Contributions to Real GDP Growth: Laspeyres Approximation less Fisher Exact**

<u>Industry Group</u>	Difference in Contribution (percentage points)		
	<u>1999</u>	<u>2000</u>	<u>2001</u>
Gross domestic product	0.00	0.00	0.00
Private industries	0.01	0.00	0.00
Private goods-producing industries	0.00	0.00	-0.00
Agriculture, forestry, and fishing	0.00	0.00	0.00
Mining	0.00	0.02	0.00
Construction	-0.01	-0.00	0.00
Manufacturing	0.01	0.01	-0.01
Durable goods	0.01	0.03	-0.01
Nondurable goods	0.00	-0.00	0.00
Private services-producing industries	0.00	0.00	0.00
Transportation and public utilities	0.01	0.01	-0.00
Transportation	0.00	0.00	-0.00
Communications	0.00	0.01	0.02
Electric, gas, and sanitary services	0.01	0.00	0.01
Wholesale trade	0.00	0.00	-0.00
Retail trade	0.00	0.01	0.00
Finance, insurance, and real estate	0.00	-0.01	-0.00
Services	-0.01	-0.01	-0.00
Government	-0.00	-0.00	-0.00
Not allocated by industry	-0.02	-0.06	0.02
Statistical discrepancy	-0.00	-0.00	-0.00
Other	-0.02	-0.06	-0.02

The differences are generally quite small for 1999-2001; all industry group differences round to less than 0.1 percentage points. The largest difference was for durable goods

<sup>11</sup> Erich Strassner of the Industry Economics Division computed the Fisher exact contributions.

manufacturing in 2000, where the Laspeyres approximation exceeded the Fisher exact contribution by 0.03 percentage points (0.92 vs. 0.89). While BEA's use of the Laspeyres approximation does not appear to have a significant impact on the computed contributions for individual industry groups, it can play a role in explaining differences between the sum of the industry group contributions and real GDP growth. For example, in 2000 the residual amount "not allocated by industry" due to factors other than the statistical discrepancy was moved farther away from zero using the Laspeyres approximation; it changed from 0.01 points using the exact formula to -0.05 points using the approximation, a difference of -0.06 points. Somewhat larger improvements would be expected for more detailed industries and for time periods with large changes in relative prices.

#### **IV. Summary and Conclusion**

This paper identifies the major sources of difference between annual estimates of real GDP growth from the NIPA's and real GDP by industry for "all industries," and provides indications of their empirical magnitudes. The difference in aggregate real output measures is important because it is one of the reasons that the published industry contributions do not sum to the growth in real GDP, clouding our understanding of how specific industries and sectors are contributing to economic growth and productivity. The principal finding of this paper is that differences in the quality, consistency, and detail of the source data—in combination with differences in methodology—are the major factor contributing to the discrepancy. The treatment of the statistical discrepancy and the specification of the contributions formula each make small contributions. Consistent source data used in a consistent conceptual framework—such as an I-O make and use tables—would result in no aggregate discrepancy, and the same contributions formula used for the NIPA's could also be used for GDP by industry.

For resolving the aggregate inconsistencies, BEA should consider and evaluate both short-run and long-run solutions. The most promising short-run option is a partial integration methodology of the kind being investigated for the upcoming comprehensive revision scheduled for release in June 2004. More consistent source data within the industry accounts—including elimination of the statistical discrepancy—should reduce aggregate real growth rate differences in most years. The Fisher exact contributions formula could then be introduced as part of this partial integration. Other short-run solutions that are feasible are not as desirable because they would



distort the relative differences in industry real growth rates. One such option is to adjust specific industry value-added quantity indexes so that the growth for “all industries” matches real GDP growth; this adjustment would be consistent with the current methodology that constrains aggregate nominal industry growth to match nominal GDP growth by including the statistical discrepancy as an “industry.” Another short-run option would be to treat the real statistical discrepancy as a balancing item for real GDP in much the same way that it is now treated on the nominal side. (Dumagan, 2002, p. 9.)

The most appealing long-run solution to the problem of inconsistent estimates is full integration of the industry and expenditures accounts using consistent source data in a consistent framework such as balanced annual I-O accounts, along with the Fisher exact contributions formula. This paper has shown that consistent data used in such a framework yields aggregate real output measures that are the same. This solution depends, however, on major improvements in the source data for gross output, final uses, and intermediate uses. Although the Census Bureau has several new initiatives designed to move towards this goal for the 2002 economic census, and the BLS continues to expand and improve service-sector producer price indexes, implementation of this solution is realistically years away. As a practical matter, then, BEA should not only continue with its efforts to integrate the industry accounts, but should also begin the process of integrating annual estimates of final expenditures in the periodic annual revisions. While this approach will not achieve the benefits of full integration, it—along with improvements in source data—will move the methodologies into closer alignment and should further reduce aggregate discrepancies.

## Appendix

From equation (13) in the main text:

$$\hat{C}_j + 1 = \frac{\sum_c [V_{cjt} - U_{cjt}] / h_{ct}}{\sum_i \sum_c [V_{cjt-1} - U_{cjt-1}] a_{ct}}$$

Recall that  $VA_{it}$  denotes nominal value added, and that  $2a_{ct} = r_c/F^P + 1$ . Let  $L_i^P$  denote the Laspeyres price index for the value added of industry  $i$ . Then we can define the constant-price industry share as:

$$\begin{aligned} \bar{s}_{it} &= \frac{\sum_c [V_{cjt-1} - U_{cjt-1}] a_{ct}}{\sum_j \sum_c [V_{cjt-1} - U_{cjt-1}] a_{ct}} \\ &= \frac{VA_{it-1}(1 + L_i^P/F^P)}{\sum_j VA_{jt-1}(1 + L_j^P/F^P)} \end{aligned}$$

Recall that  $2/h_{ct} = 1/r_{ct} + 1/F^P$ . Let  $P_i^P$  denote the Paasche price index for the value added of industry  $i$ , let  $L_i^Q$  denote the Laspeyres quantity index, and let  $P_i^Q$  denote the Paasche quantity index. For industry  $i$ :

$$\hat{C}_i + 1 = \bar{s}_{it} \frac{\sum_c [V_{cjt} - U_{cjt}] / h_{ct}}{\sum_c [V_{cjt-1} - U_{cjt-1}] a_{ct}}$$

where:

$$\begin{aligned} \frac{\sum_c [V_{cjt} - U_{cjt}] / h_{ct}}{\sum_c [V_{cjt-1} - U_{cjt-1}] a_{ct}} &= \frac{VA_{it}(1/P_i^P + 1/F^P)}{VA_{it-1}(1 + L_i^P/F^P)} \\ &= \frac{L_i^Q + L_i^P P_i^Q / F^P}{1 + L_i^P / F^P} \\ &= \frac{F^P L_i^Q + L_i^P P_i^Q}{F^P + L_i^P}. \end{aligned}$$

## References

- Commission of the European Communities, International Monetary Fund, Organization for Economic Cooperation and Development, United Nations, and the World Bank. 1993. *System of National Accounts 1993*. Brussels/Luxembourg, New York, Paris, and Washington, DC.
- Dumagan, Jesus C. *Exact Income-Side Contributions to Percent Change in GDP*. mimeo, US Department of Commerce, August 2002.
- Diewert, Erwin W. 1978. "Superlative Index Numbers and Consistency in Aggregation" *Econometrica* 46, no. 4, 883-900. Reprinted in *Essays in Index Number Theory Volume 1*, W.E. Diewert and A.O. Nakamura, eds. Amsterdam: North Holland, 1993, pp. 253-73.
- Faruqui, Umar, Wulong Gu, Mustapha Kaci, Mireille Laroch, and Jean-Pierre Maynard. 2003. "Differences in productivity growth: Canadian-U.S. business sectors, 1987-2000." *Monthly Labor Review* (April): 16-29.
- IJzeren, J. van. 1952. "Over de Plausibiliteit van Fisher's Ideale Indices" ("On the Plausibility of Fisher's Ideal Indices"), *Statistische en Econometrische Onderzoekingen* (Centraal Bureau voor de Statistiek), Nieuwe Reeks, 7, no. 2, 104-115.
- Lawson, Ann M., Kurt S. Bersani, Mahnaz Fahim-Nader, and Jiemin Guo. 2002. "Benchmark Input-Output Accounts of the United States, 1997." *Survey of Current Business* 82 (December): 19-109.
- McCahill, Robert J. and Brian C. Moyer. 2002. "Gross Domestic Product by Industry for 1999-2001." *Survey of Current Business* 82 (November): 23-41.
- Reinsdorf, Marshall B., Diewert, W. Erwin, and Ehemann, Christian. 2002. "Additive decompositions for Fisher, Törnqvist and Geometric Mean Indexes". *Journal of Economic and Social Measurement* 28: 51-61.
- Vartia, Y.O. 1976. "Ideal Log-change Index Numbers." *Scandinavian Journal of Statistics* 3: 121-6.
- Yuskavage, Robert E. 1996. "Gross Product by Industry, 1959-94." *Survey of Current Business* 82 (August): 133-155.
- Yuskavage, Robert E. 2000. *Priorities for Industry Accounts at BEA*. Paper presented at a meeting of the BEA Advisory Committee, November 17. Available on the Internet at <http://www.bea.gov/bea/papers/>