

Technical Notes

Disclosure-avoidance procedures

Like other statistical agencies, the Bureau of Economic Analysis (BEA) is legally required to safeguard the confidentiality of the information that it receives. In addition, like other agencies, it must balance its responsibility to avoid disclosing confidential information with its responsibility to release and to publish as much information as possible. It balances these responsibilities by presenting the estimates for regions, States, and local areas only at the Standard Industrial Classification (SIC) two-digit level or North American Industry Classification System (NAICS) subsector level, even though it receives source data at the SIC three- and four-digit levels or NAICS four- and five-digit industry levels.

Most of the data series that BEA receives from other agencies are not confidential. The agencies summarize this data to aggregate totals by program and by State or county, so that each record, or data cell, contains data for enough individuals or establishments to preclude the identification of the data for a specific individual or establishment and, therefore, to preclude the disclosure of confidential information.¹

However, the ES-202 tabulations that BEA receives from the Bureau of Labor Statistics (BLS) include records that would disclose confidential information. The confidential information on wages and salaries for some business firms is identifiable from the State and county estimates of wages and salaries at the SIC two-digit level or NAICS subsector level that are derived from the ES-202 data.²

1. For a list of some of the agencies that provide data to BEA, see "Sources of the data" in the "Overview."

2. For specific information, see "Wage and Salary Disbursements."

To prevent either the direct or the indirect disclosure of the confidential information, BEA uses the BLS State and county nondisclosure file.

BEA uses as many BLS nondisclosure cells as possible, but cannot use some of them for various reasons. The most important reasons are that the industry structure published by BEA does not exactly match the SIC two-digit or NAICS subsector detail provided by BLS and that BEA does not use ES-202 data for the farm sector. When BEA drops BLS nondisclosure cells, other cells must be selected to prevent the disclosure of confidential information. In order to determine which estimates should be suppressed, the total wages and salaries file and the wages-and-salaries-nondisclosure file are used to prepare a multidimensional matrix. This matrix is tested, and the estimates that should be suppressed are selected.³

Dual allocation

The dual allocation procedure is used to allocate a data series by State and by industry simultaneously to a set of national control totals for the industries and to a set of all-industry control totals for the States. It is also used to allocate a data series by county and by industry simultaneously to a set of State control totals for the industries and to a set of all-industry control totals for the counties.

In a dual allocation, the two sets of control totals are placed in a matrix, with the national (or State) control totals as the column totals and the State (or county) control totals as the row totals. The allocating series is placed in the same matrix to serve as the set of elements. These elements are adjusted alternately by allocation to sum first to the column totals and then to the

3. In this test, computer programs impose a set of rules and priorities on this matrix so that the estimates that should be suppressed are selected until indirect disclosure is impossible.

row totals; the procedure is completed by an allocation of the column totals.

Employment

The BEA employment series for States and local areas comprises estimates of the number of jobs, full-time plus part time, by place of work. Full-time and part-time jobs are counted at equal weight. Both employment for wages and salaries and proprietors' employment are included, but the employment of unpaid family workers and volunteers is not included. Proprietors' employment consists of the number of sole proprietorships and the number of partners in partnerships. The description "by place of work" applies to the wage and salary portion of the series and, with relatively little error, to the entire series. The proprietors' employment portion of the series, however, is more nearly by place of residence because, for nonfarm sole proprietorships, the estimates are based on IRS tax data that reflect the address from which the proprietor's individual tax return is filed, which is usually the proprietor's residence. The nonfarm partnership portion of the proprietors' employment series reflects the tax-filing address of the partnership, which may be either the residence of one of the partners or the business address of the partnership.

The employment estimates are designed to be consistent with the estimates of wage and salary disbursements and proprietors' income that are part of the personal income series. The employment estimates are based on the same sets of source data as the corresponding earnings estimates and are prepared with parallel methodologies. Two forms of proprietors' income--the income of limited partnerships and the income of tax-exempt cooperatives--have no corresponding employment estimates.

Imputation

Imputations are added to personal income and to other measures in the national income and product accounts (NIPA's) so that a comprehensive account of total production and its distribution can be presented. The imputed transactions included in the NIPA's are a limited set of exceptions to the principle that the NIPA's reflect market transactions in goods and services. In order to keep the NIPA measures invariant to how certain activities are carried out, imputations are made to place a market value on certain transactions that do not occur or are not observable in the market economy. In addition, some market transactions are reconstructed to provide a representation of the activity that is more appropriate for the NIPA's. Both a measure of the production and the incomes associated with that production are imputed. The imputations described here are those that affect personal income.⁴

Specifically, six imputations are included in the estimates of personal income: Imputed pay-in-kind, employer-paid health and life insurance premiums, the net rental value of owner-occupied farms and the value of food and fuel produced and consumed on farms, the net rental value of owner-occupied nonfarm housing, the net margins on owner-built housing, and the imputed interest paid by financial intermediaries except life insurance carriers. These imputations accounted for about 8 percent of personal income at the national level in 1999.

Imputed pay-in-kind is added to the estimates of wages and salaries so that all the earnings of employees who receive part of their wages in pay-in-kind will be included in personal income. This imputation is an estimate of the value of the food, lodging, clothing, and other goods and services that are received by employees from their employers as full payment or as partial

4. See table 8.21, "Imputations in the National Income and Product Accounts," Survey 82 (August 2002): 118.

May 6, 2003

payment for their services. For additional information, see "Pay-in-kind" in the "Glossary."

The net rental value of owner-occupied farm housing and the value of food and fuel produced and consumed on farms are included in farm proprietors' income so that that measure reflects the income from all of the production of noncorporate farms.

The net rental value of owner-occupied nonfarm housing is included in the rental income of persons. The imputation assumes that the owner-occupants are in the rental business and that they are renting the houses in which they live to themselves: As tenants, they pay rent to the landlords (that is, to themselves); as landlords, they collect rent from their tenants (that is, from themselves), they incur expenses, and they may have a profit or a loss from the rental business.

The net margins on owner-built housing is included in proprietors' income, classified in the construction industry. It is the imputed net income of individuals from the management of the construction or renovation of their own dwellings and is included in the measure of the output of structures.

The imputed interest income paid by financial intermediaries except life insurance carriers, which is included in personal interest income, is received by persons from depository institutions, that is, from commercial banks, mutual savings banks, savings and loan associations, credit unions, and regulated investment companies. It is an estimate of the value of the services (such as checking and record keeping) that these institutions provide to persons without an explicit charge.

Another portion of personal interest income is classified as imputed interest. This is an estimate of the investment income that is earned on the financial reserves of life insurance carriers. This investment income is attributed to the policyholders in order to include it in personal saving, and not in business saving, and when the income is earned, rather than when it

is distributed.

Industry classification

For the private sector, the Standard Industrial Classification (SIC) provided by the Office of Management and Budget is used for the classification of the estimates of wage and salary disbursements, other labor income, and proprietors' income by establishment industry. The Standard Industrial Classification Manual, 1967 is used for the classification of the estimates for 1969-74, the 1972 Manual is used for the classification of the estimates for 1975-87, and the 1987 Manual is used for the classification of the estimates for 1988-2000.⁵

The North American Industry Classification System (NAICS) is used for the classification of the estimates for 2001.⁶

For the public sector, the estimates of wages and salaries and other labor income are classified by level of government--Federal, State, and local. The estimates for the Federal Government are subclassified into civilian and military.

5. Executive Office of the President, Office of Management and Budget, Statistical Policy Division, Standard Industrial Classification Manual, 1967 (Washington, DC: U.S. Government Printing Office (GPO), 1967); Manual, 1972 (GPO, 1972); Manual, 1987 (GPO, 1987).

6. Office of Management and Budget: North American Industry Classification System, United States, 2002 (Lanham, MD: Bernan Press, 2002)

Interpolation and extrapolation

Interpolation and extrapolation are used to prepare the first approximations of the State estimates of some of the components of personal income for the years in which direct source data are unavailable. Both procedures use the data for these components for benchmark years--the years for which the best data are available--and both frequently use other data that are related to the benchmark-year data for the components.

Interpolation is used to derive the first approximation of the estimates for years that are between benchmark years. For example, if data for wages and salaries for an industry were available only from the decennial censuses of population but employment data were available annually from another source, the first approximations of wages and salaries for 1981-89 are interpolated from the State data for wages and salaries 1980 and for 1990, the two census benchmark years, and from the data for employment for 1980-90.

Extrapolation is used to derive the first approximations for the years that are beyond the most recent benchmark year. For example, the first approximations of wages for 1991-99 might be extrapolated from the census benchmark data for 1990 and from the employment data for 1990-99. The estimates based on extrapolation are usually superseded by revised estimates when benchmark data become available for a more current year. For the preceding example, the estimates for 1991-99 would be superseded by estimates based on interpolation when census benchmark data became available for 2000.

Both interpolation and extrapolation are illustrated in the following examples. In the first two examples, interpolation is used to derive the first approximations of wages and salaries for an industry in counties A, B, and C for the years 2 and 3 that are between the benchmark years 1 and 4. In the third example, extrapolation is used to derive the approximations for year 5.

In the first example, "straight-line interpolation" is used to derive the first approximations for years 2 and 3 from the State data for the benchmark years 1 and 4.⁷ The first approximations for year 2 equals the amount for year 1 plus one-third of the increase from year 1 to year 4; the preliminary estimate for year 3 equals the amount for year 1 plus two-thirds of the increase.

Wages and salaries in thousands of dollars

	Year 1 (benchmark)	Year 2 (interpolation)	Year 3 (interpolation)	Year 4 (benchmark)
County A	28	34	40	46
County B	34	43	53	62
County C	74	81	87	94

In the second example, interpolation with a related series of data, the indicator series, is used to derive the first approximations for years 2 and 3 from the benchmark data for years 1 and 4 and from the indicator series for all 4 years. The data for wages and salaries are the benchmark data, the employment data are the indicator series, and the average wages (computed as wages and salaries divided by employment) are the interpolation ratios.⁸ This method of

7. Straight-line interpolation assumes that the magnitude of the annual change is the same in each year in the interpolated time series, subject to modification by the adjustment to the State control totals. Straight-line interpolation is used as the default option, when no annual source data related to the income series are available.

8. Using an indicator series for interpolation between 2 benchmark years assumes that any change in the relationship between the data for the income component for the benchmark years and the data from the indicator series for the benchmark years occurs uniformly over time. This relationship is embodied in the interpolation ratios, which in this example are the average wages. For this procedure, straight-line interpolation of the benchmark-year interpolation ratios is used to calculate the ratios for the intervening years. A benchmark-year interpolation ratio is the ratio of the datum for an income component for the benchmark year to the datum for the same year from the annual indicator series. The interpolation ratios for the intervening years are multiplied by the data for those years from the indicator series to yield the interpolated series for those years.

interpolation is illustrated in three steps.

First, the average wages of the employees in an industry for years 1 and 4 are calculated from data for wages and salaries and data for employment for those years. The wages for each year are divided by the number of employees for the year to yield the average wages of the employees.

Employment and average wages in dollars

	<u>Year 1</u>		<u>Year 4</u>	
	Employment	Average wages in dollars	Employment	Average wages in dollars
County A	4	7,000	4	11,500
County B	6	5,667	10	6,200
County C	11	6,727	10	9,400

Second, straight-line interpolation is used to derive the average wages for years 2 and 3 from the average wages for years 1 and 4.

Average wages in dollars

	<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>	<u>Year 4</u>
	(benchmark)	(interpolation)	(interpolation)	(benchmark)
County A	7,000	8,500	10,000	11,500
County B	5,667	5,845	6,022	6,200
County C	6,727	7,618	8,509	9,400

Third, the interpolated average wages for each year are multiplied by the employment data for each year to yield the first approximations.

Employment and wage approximations

	<u>Year 2</u>		<u>Year 3</u>	
	<u>Employment</u>	<u>Wages in thousands of dollars</u>	<u>Employment</u>	<u>Wages in thousands of dollars</u>
County A	5	43	4	40
County B	7	41	9	54
County C	10	76	9	77

In the third example, extrapolation with an indicator series is used to derive the first approximations of wages for year 5 from the average wages for year 4--used here as the extrapolation ratios--and employment data for year 5.⁹ The average wages are multiplied by the employment data to yield the first approximations of wages for year 5.

First approximations of wages for year 5

	<u>Year 4</u>	<u>Year 5</u>	
	<u>Average wages in dollars</u>	<u>Employment</u>	<u>Wages in thousands of dollars</u>
County A	11,500	5	58
County B	6,200	12	74
County C	9,400	9	85

After interpolation or extrapolation is used to calculate the first approximations of a component of State personal income, the approximations are adjusted by allocation to sum to the national estimate of the component.

Per capita personal income

This measure of income is calculated as the personal income of the residents of a given area

9. Using an indicator series for extrapolation assumes that the relationship between the data for the income component for the latest benchmark year and the data from the indicator series for that year remains unchanged in the subsequent years.

May 6, 2003

divided by the resident population of the area. In computing per capita personal income for States and counties, BEA uses the Census Bureau's annual midyear population estimates. Except for the college student and other seasonal populations, which are measured on April 1, the population for all years is estimated on July 1.

The local area per capita personal income estimates should be used with caution for several reasons. In some instances, an unusually high or low per capita personal income is the temporary result of unusual conditions, such as a bumper crop or hurricane. In other instances, the income levels of certain groups atypical of the resident population may cause a longer term high or low per capita personal income that is not indicative of the economic well-being of the area. For instance, a major construction project--such as a defense facility, power plant, or dam--may substantially raise the per capita personal income of an area for several years because it attracts highly paid workers whose income is measured at the construction site. This high per capita income is not indicative of the economic well-being of most of the residents of the area (or, in many cases, of the resident construction workers themselves, because they frequently send a substantial portion of their wages to their dependents living in other areas).

Conversely, the presence of a large institutional population--such as that of a college or a prison--will tend to keep the per capita personal income of an area at a lower level because the residents of these institutions have little income attributable to them at these institutions. This lower per capita personal income is not indicative of the economic well-being of most of the residents of the area (or, in some cases, of the institutional populations, because some of these populations, such as college students, typically receive support from their families living in other areas).

The per capita personal income estimates can also be misleading in areas where population

changes rapidly. Population is measured at midyear, whereas income is measured as a flow over the year; therefore, a significant change in the population of an area during the year, particularly if it occurs around midyear, can cause a distortion in the per capita personal income estimates.

In counties where farm income predominates, additional considerations should be taken into account. Farm proprietors' income as measured for personal income reflects returns from current production; it does not measure current cash flows. Sales out of inventories are included in current gross farm income, but they are excluded from net farm income because they represent income from a previous year's production. Additions to inventories are included in net farm income at current market prices; therefore, farmers' attempts to regulate their cash flows by adjusting inventories are not reflected in BEA's farm proprietors' income estimates. However, this regulation of cash flows by farmers extends their earnings cycles, so it helps them to survive losses or lowered income for 2 or 3 years. In addition, the per capita personal income of sparsely populated counties that are dependent on farming will react more sharply to vagarious weather, world market demand, and changing government policies related to agriculture than that of counties where the sources of income are more diversified.

Personal income, adjusted gross income, and money income

The measure of personal income that is prepared by BEA differs substantially from adjusted gross income (AGI), which is the principal measure of the income of individuals that is tabulated by the Internal Revenue Service. Personal income also differs from money income, which is prepared by the Census Bureau.

As compared with AGI, personal income consists of the income of nonprofit institutions

serving individuals, private noninsured welfare funds, and private trust funds as well as of individuals, whereas AGI consists only of the income of individuals who file individual income tax returns. Personal income also includes employer contributions to private health and pension funds and to government employee retirement plans, several types of imputed incomes, transfer payments, and all of the interest received by individuals, whereas AGI excludes all employer contributions, imputed incomes, most transfer payments, and the nontaxable interest received by individuals. Personal income, unlike AGI, excludes personal contributions for social insurance, realized capital gains and losses, and pensions and annuities from private and government employee retirement plans.¹⁰

Personal income differs from money income mainly because money income consists only of the income that is received by individuals in cash and its equivalents. Personal income, unlike money income, includes imputed income, lump-sum payments not received as part of earnings, certain in-kind transfer payments--such as medicaid, medicare, and food stamps--and employer contributions to private health and pension funds and to government employee retirement plans. Personal income, unlike money income, excludes personal contributions for social insurance, pensions and annuities from private and government employee retirement plans, and income from interpersonal transfers, such as child support.

The State personal income measure for a year includes the income received by individuals living in the State during the year. The State money income measure for a year consists of the income received during the year by the individuals living in the State on April 1 of the following year, regardless of where they were living when they received the income. The income received

10. For more information, see Thae S. Park, "Comparison of BEA Estimates of Personal Income and IRS Estimates of Adjusted Gross Income," Survey 82 (November 2002): 13-20.

May 20, 2003

by individuals who died or moved abroad before April 1 of the following year is not included in the money income measure of any State.

Personal income for States is prepared quarterly and annually and for counties is prepared annually, whereas money income for States, counties, and cities is prepared decennially from the data from the "long-form" sample conducted as part of the census of population.¹¹

11. The most recent estimates of money income for States and counties--those for 1999--were prepared from data from the 2000 Census of Population. In addition, the Census Bureau has prepared estimates of median household income for States and counties, most recently through 1999.

For the Nation and for the four census regions, the Census Bureau also prepares annual estimates of money income from the data from the current population survey.