

IOWA STATE UNIVERSITY

**Data in Context: Notes on Employment and Labor Force
Data**

Mark D. Imerman

April 2006

Working Paper # 06018

**Department of Economics
Working Papers Series**

Ames, Iowa 50011

Iowa State University does not discriminate on the basis of race, color, age, religion, national origin, sexual orientation, gender identity, sex, marital status, disability, or status as a U.S. veteran. Inquiries can be directed to the Director of Equal Opportunity and Diversity, 3680 Beardshear Hall, (515) 294-7612.

Data in Context:
Notes on Employment
And
Labor Force Data

Mark Imerman

Department of Economics
Iowa State University

April 24, 2006

Table of Contents

<u>Confidentiality, Disclosure, and Access</u>	1
<u>Employment Rates, Unemployment Rates, and Resident Labor Force</u>	2
Resident labor force	
Employment	
Place of residence	
Major data source: Current Population Survey (CPS)	
Major data source: Local Area Unemployment Statistics (LAUS)	
Major data source: Decennial Census	
<u>Employment by Industry</u>	5
Industrial classification	
Ownership, industrial classification, and the National Income and Product Accounts (NIPA)	
"Job-based" rather than "People-based" statistics	
Place of work	
Major data source: State Labor Market Information (LMI)	
Major data source: U.S. Bureau of Economic Analysis (BEA)	
Major data source: Current Employment Statistics (CES)	
Major data source: U.S. Bureau of the Census	
Employment by industry by place of residence, data from the decennial Census	
<u>Employment by Occupation</u>	10
Major data source: Decennial Census	
Major data source: Occupation Employment Statistics (OES)	
Major data source: Iowa Department of Workforce Development LMI program	

This paper is also available in HTML format at

http://www.econ.iastate.edu/prosci/imerman/mypapers/DC_empdata.html

Data in Context: Notes on Employment and Labor Force Data

Employment is the commitment of personal effort to the process of producing goods and services. In common usage, a person is said to be "Employed" if the person works for another person, a person is said to be "Self-employed" if the person works for him or herself, and a person is said to be "Unemployed" if that person has no job. This terminology works fine in most casual conversations, but breaks down when dealing with data collection and maintenance efforts. In order to maintain more precision and consistency, these terms are more tightly controlled in the official statistical process. Understanding the terms and the sources of data are important to understanding and correctly using the information that the data provide. The discussion below will provide definitions of terms and an understanding of data collection and content for major statistics on "Employment" in its three most common contexts:

- Employment and unemployment rates for the residential workforce
- Employment by industry
- Employment by occupation

There may be several sets of employment statistics available for any specified area. These might originate from multiple statistical agencies and refer to different measures of "Employment." It is important to understand what the statistics are and what they tell us.

Confidentiality, Disclosure, and Access

In addition to knowing what the statistics tell us, it is important to realize what the statistics cannot tell us. Most of the major employment-related statistics are collected under a legal mandate. This means that individuals and businesses cannot refuse to supply the information required by the government. These mandated programs include payroll tax reporting, income tax reporting, responses to census and other special surveys, etc. To assure that responses to these mandates are honest and complete, the federal and state governments, in nearly all cases, guarantee confidentiality to the respondents. This means that in cases where data describing a specific individual or business entity would be recognizable in the data set, that portion of the data set is suppressed. It is not disclosed or made public. This is why employment and earnings data is available on only a very limited or highly aggregated basis for local areas.

In general, when data is suppressed, some sort of disclosure code is presented in its place. This disclosure code will tell the data user why this data is not being made available. The three most common reasons for suppressing data items are

- Confidentiality considerations - Generally, when there are three or fewer respondents in a category or when one of the respondents is relatively large when compared to the other respondents, as a group, confidentiality rules (which are also usually mandated by law) require that the data item be suppressed.
- Statistical confidence - Where statistical estimation is used and the sample size for a data item is too small to generate reliable statistical results, the data is generally suppressed.
- Incomplete data - When the data collection process has been interfered with and is not complete for individual data items or geographic areas, those data items may be suppressed.

For any individual data set, there may be other specific reasons for data suppression. It is important to know if and why data is suppressed. This information is invaluable in determining where and when gaps in the data might be treated as a numeric zero (a data subject exists but was inactive for the period), as null (no data subject exists), or as suppressed (reportable activity exists but is temporarily or permanently unavailable).

Finally, it is a common misconception to look at data suppressions as weak spots in the data collection and distribution process and as unnecessary impositions upon the data user. It would be more correct to look at the consistent application of a suppression system as a major source of database strength. Without guaranteeing respondent confidentiality and privacy, we could not take for granted the pervasive existence of an honest reporting ethic. Consistent suppression and the maintenance of confidentiality are primary reasons that we can have confidence in the data that is made available.

Employment Rates, Unemployment Rates, and the Resident Labor Force

Statistics from the U.S. Bureau of Labor Statistics (www.bls.gov), the U.S. Census (www.census.gov), the Iowa Department of Workforce Development (www.iowaworkforce.org) (or any of the state agencies charged with administering employment security programs) that refer to “Employment” rates or “Unemployment” rates for a given area (state, county, or metropolitan area) are estimates of the percent of the area’s resident labor force that is employed or unemployed. For any area, the sum of these two numbers will equal one hundred percent. If a person is a member of the “Labor force,” that person is either employed or not employed and cannot be both. If you have an area’s employment rate, you can subtract it from 100 percent to get the unemployment rate, and vice versa.

Resident labor force

An area's resident labor force consists all members of the civilian noninstitutionalized population aged 16 or over that are either employed or are actively seeking work. For the purposes of employment statistics, a person is employed if that person is personally engaged in any productive activities for which compensation is received. Regardless of

whether a person is working part-time, full-time, or over-time, the employment rate statistics will treat that person the same. Members of the labor force that are compensated for any level of employment are counted as "Employed," regardless of how adequate or appropriate that employment might appear. Members of the labor force that are not counted as "Employed" are counted as "Unemployed."

Employment

Employment refers to personal productive effort that is compensated for regardless of whether the person employed works for and is paid by another or works for and is paid by his or her own business. Individuals that work for others are generally referred to as "Wage and salary" employees in the statistics. Individuals that work for themselves or for their own business are generally referred to as "Proprietors." Statistics that simply refer to "Employment" generally include both proprietors and wage and salary workers.

Proprietors can either be "Sole proprietors," which individually own their own business or "Partners," in which they directly share ownership of a business with others. The term, "Proprietor," does not refer to a person that owns stock in a corporate business. Compensation for stock ownership is the return on providing capital rather than compensation for personal effort or labor to the productive process. This compensation for providing capital is reported as "Dividends" which, like earnings from employment, are a component of total personal income, but they are not attributable to personal productive effort.

Place of residence

Estimates of employment rates and unemployment rates for a given area are "Place of Residence" statistics. These estimates refer to people living within a designated area, regardless of where those people work. For example, a county might have a very high employment rate (measure of residents engaged in the labor force that have a job) even if there are a relatively small number of jobs in the county if labor force participants are able to commute to jobs in surrounding counties. On the other hand, there could be a relatively large number of jobs in a county with a high unemployment rate if individuals that commute in from other counties hold many of the jobs that are available.

Major data source: Current Population Survey (CPS)

The official nationwide estimates of the size of the labor force, employment rate, and unemployment rate are obtained from the CPS. The CPS is a joint project of the U.S. bureaus of the census and labor statistics. CPS conducts a monthly telephone survey of 60,000 households, nationwide, to determine labor force and employment status. Monthly data available from the CPS include nation-wide estimates of labor force participation, employment and unemployment by industry, educational attainment (for population 25 and over), age, ethnicity, race, employee class, discouraged workers, and people holding multiple jobs. Data is available from (www.bls.gov/cps/cpsatabs.htm). The CPS also provides annual estimates of state-level labor force and employment.

The strengths of the CPS are that the sample and the collection process are consistently controlled, the data is timely (data for the most recent previous month is generally available), and connects many demographic and motivational characteristics to the labor force participation data. The weakness of this source is that the sample size is insufficient to provide local data.

Major data source: Local Area Unemployment Statistics (LAUS)

The Local Area Unemployment Statistics (LAUS) program provides estimates of total nonfarm labor force and employment for states and metropolitan areas at data.bls.gov/cgi-bin/surveymost?la through the BLS, and for states, counties, and urban areas with populations over 25,000 through the individual state's Labor Market Information programs (for Iowa data, go to www.iowaworkforce.org/lmi/laborforce/index.html).

This is a cooperative project between the U.S. Bureau of Labor Statistics and the individual state's employment security agencies. LAUS estimates these statistics using data from the CPS, current employment statistics, state unemployment insurance program statistics, and other sources. The major strengths of this data source are its timeliness and local level. The major weaknesses are that LAUS does not include farm employment and does not tie employment statistics to any demographic information.

Major data source: Decennial Census

The decennial Census (www.census.gov/) also provides information on the size of the potential labor force, labor force participation, and employment or unemployment. This data is available for smaller geographic areas than either the CPS or LAUS statistics. Labor force statistics from the Census are estimated from responses to the decennial Census long form, which was distributed to approximately one-sixth of U.S. households with the 2000 Census. This is, by far, the largest sampling used in the development of labor force data. The size of this sample is the source of both the strengths and the weaknesses of the data set.

The drawbacks of this data are timeliness and self-reporting. Decennial Census data is available only on ten-year intervals and estimates are not publicly available until almost two years after survey collection. Also, collection processes for decennial Census data do not include the strict control that the CPS and major source data for the LAUS data sources receive.

On the plus side, decennial Census data is the only source of labor force and employment data that links employment rate estimation, employment by industry estimates, and employment by occupation estimates to the same source-data collection process. Census data can also be directly linked to several demographic variables, including household characteristics of the labor force. Finally, Census data is consistently available for much smaller geographic areas than are any other sources of labor-force data. All of these characteristics make Census labor-force data a useful bridge when information from

multiple sources is tied together to track labor-force and employment trends and conditions between Census years.

Basic labor force data from the 2000 Census is distributed as part of Standard File 3 (SF3) tables P43 through P48.

Employment by Industry

Local statistics from the U.S. Bureau of Economic Analysis (<http://www.bea.gov/>), the U.S. Bureau of Labor Statistics (www.bls.gov/), or the agency in any state charged with administering employment security programs (in Iowa, the Iowa Department of Workforce Development (www.iowaworkforce.org/)), that refer to employment by industry for a given area (state, county, metropolitan area, etc.) refer to the number of jobs reported by employers within the area. These statistics are based on the payroll tax reports that employers file in compliance with state and national employment security legislation (unemployment insurance, workers' compensation, etc.)

Industrial classification

For reporting purposes, every covered employer (independent employment locations mandated to participate in federal employment security programs) is placed in a single industrial classification (www.census.gov/epcd/www/naics.html). All jobs reported by that employer are counted as employment within that industry, regardless of the specific task assigned to any of the individual jobs. For example, a shuttle-bus driver that delivers workers from remote parking areas to their jobs in a telemarketing center would be classified as working in the telemarketing industry (NAICS 561422) rather than in the bus transportation industry (NAICS 4854). The opposite would be true of a telephone marketer working for a company that contracts chartered busses.

Ownership, industrial classification, and the National Income and Product Accounts (NIPA)

In addition to being classified by industry, every covered employer is also classified into one of four employer ownership categories:

- Federal government
- State government
- Local government
- Private

This allows all reported job counts and payroll information to be identified by industry and by type of enterprise (public or private). As the data moves through the state and federal data systems to support a variety of administrative purposes, different agencies report statistics in different ways with regard to ownership and industry.

Data from the U.S. Bureau of Labor Statistics (BLS) and the state employment security agencies are generally reported on a strict industrial-classification basis. Data are reported by the type of industry that most closely matches activities at that employment site, regardless of whether that activity is being undertaken by a public or a private enterprise. This means that the employment and payroll of power plants, airports, etc. that are government owned but managed as separable entities are reported within the transportation and public utilities industries rather than as part of the government administration industry.

On the other hand, data from the U.S. Bureau of Economic Analysis (BEA) is reported on a NIPA basis. NIPA separates all public employment from private employment as an initial step in the national accounting process. This means that all public employment and payroll are reported as government employment and payroll by the BEA, regardless of industry. BEA data is useful if we are looking for the total of government jobs and payroll. BLS data comes closer to the mark if we are looking only for the number of public administration employees and are not interested in nonadministrative public enterprise employment.

Finally, under both systems of reporting, employees and payroll of private firms that perform public administrative duties through contracts or other arrangements are never reported under the public administration industrial classifications. This is because the public administration classifications are defined in such a way that private employers are excluded from these classifications regardless of primary productive activity.

"Job-based" rather than "People-based" statistics

It is important to note that these statistics are primarily based on "Jobs" rather than "People." The statistics enumerate or estimate (depending upon the source used) the number of jobs filled (for which earnings are reported) within any given area. Jobs are counted the same whether they are filled on a full-time, a part-time, or an over-time basis. Jobs are counted regardless of whether they are the primary source of income and regardless of whether the jobholder is satisfied with the employment. For example, if an individual holds three part-time jobs and works a total of 27 hours per week, the statistics report three jobs rather than one individual. Similarly, if one individual works at a single job for 60 hours per week, the statistics report one job. In the case where two individuals work out a job sharing agreement with an employer to split a single 40-hour-per-week job, the statistics will show two jobs, because the employer has to report payroll taxes for both individuals as if they had separate jobs.

Place of work

Employment by industry data provided by state employment security agencies and the U.S. bureaus of labor statistics and economic analysis are reported by "Place of work." This means that jobs are reported where they exist, regardless of where the employee resides. If an employer in Area A provides a job that is held by a resident of Area B, the employment by industry statistics will count the job-by-industry in Area A, where the job

exists. Recall from the discussion about residential labor force, above, that the fact that a person residing in Area B holds a job in Area A shows up in the employment rate, the unemployment rate, and the labor-force participation statistics of Area B. The employment by industry data are reported by place of work, the labor force and employment (unemployment) rate data are reported by place of residence. Most employment by industry data are "Jobs" based. Labor force and employment-rate data are "People" based.

Major data source: State Labor Market Information (LMI)

Employment by industry data from the individual states' employment security agencies, the U.S. Bureau of Labor Statistics, and the U.S. Bureau of Economic Analysis are based on employer payroll tax reports collected by the individual states' employment security agencies. The most current of these sources is generally the LMI system, maintained by the states' employment security agencies. Most state LMI systems supply monthly enumerations of employment by industry for individual counties and places with populations of over 25,000 people. This data is very current, with monthly data generally available with only a one or two month delay. A list of contacts for individual state LMIs is available at www.bls.gov/bls/ofolist.htm. Iowa data is available at www.iowaworkforce.org/lmi/laborforce/index.html - counties.

Employment by industry data reported by the states' LMI are a strict enumeration of the data received on payroll tax reports. This data is generally labeled as "Nonfarm covered employment" data. This means that no attempt is made to estimate employment in businesses that are not subject to employment security regulations and that no statistics are released for the farm industry as a whole (because most of the farm industry is exempt from employment security regulations). The strength of this data set is its timeliness and its local content. Its major weakness is that there is no accounting for employment that is not covered by employment security regulations.

Major data source: U.S. Bureau of Economic Analysis (BEA)

County-level employment and payroll data available from the Bureau of Economic Analysis are generally available in May of the second year following the calendar year that the estimated data describes. The BEA data are estimates rather than enumerations, and the statistics are regularly re-estimated over the first three years that they are available. This allows updated information to be incorporated into the estimates as it becomes available. Also, entire BEA data series are re-estimated at any time that significant improvements are made to the estimation process.

Regular re-estimation requires that BEA data be regularly replaced rather than simply added to. Re-estimation also means that the historical data set available from the BEA is nearly always compatible with the current estimate. This is not always the case with data derived from administrative sources.

While BEA estimates are based on employment security data, the BEA also uses many other sources of data in their estimation process (Census demographics, tax reporting, public assistance participation, agricultural programs, etc.) This gives the BEA the ability to include estimates of employment and payroll not covered by employment security programs and to estimate noncash compensation. This makes the BEA estimates more inclusive than the other public sources of employment by industry data, and is a major strength of this data set. The trade-off is that the BEA county-level data are available only after an 18-month lag and are open to major revisions for an additional two years.

Major data source: Current Employment Statistics (CES)

In the continuum from strict enumeration (state-level data) and full estimation (BEA data), employment by industry data from the U.S. Bureau of Labor Statistics Current Employment Statistics (CES) program fall somewhere in the middle. CES employment data is estimated from a sample of approximately 300,000 businesses (employing about 1/3 of all payroll workers), nationwide, on a monthly basis www.bls.gov/sae/home.htm.

CES provides estimates of employment by industry and average weekly hours and earnings by both industry and employee class. Data are available for states and metropolitan areas. The strength of this data is that it is timely and it is consistently produced within any state and at the federal level. A weakness of the data is that it does not include complete farm or nonpayroll employment. Also, while estimates are constructed to be consistent with data available within any individual state, the individual state estimates are affected by sampling errors that differ on a state-by-state basis and between state and national estimates. For this reason, the BLS issues a special caution on the aggregation of these data on a regional basis.

Major data source: U.S. Bureau of the Census

The U.S. Bureau of the Census puts out a number of data compilations related to employment by industry by place of work. The three most commonly used are the Economic Census, County Business Patterns, and The County and City Data Book.

The Economic Census is a mandatory survey of most major businesses in the United States. The survey is done every five years. Data is available for surveys taken through 2002. The Economic Census includes data by industry and area on number of firms responding (categorized by whether they have a payroll or not), total employment, total wage costs, and total payroll costs (wage plus nonwage employment costs). This data is available at a state, county, community, and zip code basis for many industries. A list of geographic coverage by industry is available at www.census.gov/epcd/ec02/guide.html.

County Business Patterns is released annually, and provides information by major industry on the number of firms, employment for the week of March 12, annually, and payroll for states and counties. Local data is based, primarily, on the payroll reporting data that comes through state employment security agencies and the U.S. Bureau of Labor Statistics. County Business Patterns excludes data on self-employed individuals,

employees of private households, railroad employees, agricultural production employees, and most government employees. The data that is published in County Business Patterns is also benchmarked to national survey data from the Economic Census, the "Annual Company Organization Survey," the "Annual Survey of Manufactures," and other sources. This process results in some minor differences between local employment and payroll data from County Business Patterns (after accounting for exclusions) and local employment and payroll data from the BLS. Information on accessing County Business Patterns is available from the Census Bureau at www.census.gov/epcd/cbp/view/cbpview.html.

The County and City Data Book provides a compilation of data, employment and otherwise, from a wide variety of sources. It is published annually, and is a handy single-source reference that is available online, on cd-rom, and in printed form. More information on this reference is available from the Census at www.census.gov/statab/www/ccdb.html.

Employment by industry by place of residence, data from the decennial Census

The decennial Census also provides data on employment by industry data. In contrast to the data that is compiled or estimated from employment security payroll taxes, the Census data is estimated from responses to the long form of the decennial Census and is available as part of file SF-3 of the U.S. Census data system. There are significant differences between this data and the data that is available from the BLS, BEA, and state employment security agencies.

- First, Census data is self-reported. In contrast to the employment security system, where trained classifications staff classifies each covered business, Census respondents choose their classifications, themselves. This leads to problems in distinction similar to the telemarketing-transportation example discussed above. There are no fixed rules enforced upon Census respondents to assure that this sort of situation results in consistent reporting.
- Second, Census data estimates the people employed by industry rather than the number of jobs filled within an industry. Census employment-by-industry data are "People" based, as opposed to "Jobs" based.
- Third, Census employment-by-industry data is reported by place of residence rather than by place of work. This means that the Census data for an area includes all of the residents that live within that area regardless of where they work.
- Fourth, Census data is available on a sub-county (township, place, census area) geography. This provides much more geographic specificity in the Census data than that available through the employment security reports.
- Finally, decennial Census data are available only for Census years, which are the final year of every decade (1980, 1990, 2000, etc.)

Employment by industry data from the 2000 Census is distributed as part of Standard File 3 (SF3) table P49.

Employment by Occupation

Statistics of employment classified by occupation are significantly different than statistics that classify employment by industry. "Occupation" refers to the type of task and skill level of an individual worker's position (secretary, machinist, teacher, truck driver, doctor, cook, etc.) "Industry" refers to the primary type of production done by the employer of that individual worker (manufacturing, retail trade, construction, mining, banking, etc.) A secretary (occupation), for example, might be employed in a manufacturing firm, a doctor's office, a university, or a farm production enterprise, all of which would be in different industries. The same would be true for almost any occupation. Official occupation classifications are available at www.bls.gov/soc/.

Occupational employment data is generally survey based, because most payroll and employment reporting is done by industry rather than by occupation. Employment by occupation statistics are available from the decennial Census, the U.S. Bureau of Labor Statistics, the Iowa Department of Workforce Development, and many of the individual states' employment security agencies.

Major data source: Decennial Census

The most comprehensive of these surveys is the decennial census of the United States. Occupational data from that source is estimated for the population as a whole on the basis of responses to the Census long form, which is filled out by approximately 16.67 percent of U.S. households.

Employment by occupation data from the decennial census is estimated by place of residence. This means that the estimates include everyone that lives in the defined area, regardless of where they work.

This dataset has three major strengths:

- The size of the sample used in estimation is quite large, giving confidence that the estimations accurately reflect the responses that we would expect from the population as a whole
- The results are available for very small geographic areas. This is possible because of the relatively large size of the sample
- The reported results are consistent across the entire U.S., making it easy to do comparative analysis.

There are also two significant weaknesses in this data source. The first is data timeliness. The data is only collected on ten-year intervals, and the responses are nearly two years old before local estimation results are released. The second is that the data is self-reported and data collection is largely unsupervised. This is not as great a problem with occupational responses as it is with industrial classification, but it still has the potential to reduce the accuracy of the estimates.

Basic labor force and employment data from the 2000 Census is distributed as part of Standard File 3 (SF3) tables P43 through P48.

Major data source: Occupation Employment Statistics (OES)

While the decennial Census is the most comprehensive of these surveys, the most consistently controlled is the U.S. Bureau of Labor Statistic's Occupation Employment Statistics program. This program surveys employers during the last quarter of the calendar year to estimate state and national statistics on employees and wages by occupational class. Estimates are usually available about one year after the survey is completed.

This data is collected from a survey of employers. As such, the data is presented on a "Place of work" basis. Data is available at www.bls.gov/oes/home.htm.

Major data source: Iowa Department of Workforce Development LMI program

In Iowa, estimates of occupational employment and wages can be obtained for the most recent year through the state's Labor Market Information program. In 2002, these estimates became available at the county level. Prior to 2002, they were only available for multicounty regions. These data are based upon the most recent OES survey from the BLS in conjunction with estimates based on locally available data. While the Iowa Department of Workforce Development indicates that these local estimates are not completely compatible with national and state OES survey results, the estimates do give a level of local detail that can be very helpful in community analysis. Data can be obtained at www.iowaworkforce.org/lmi/occupations/index.html.

Good luck.

Mark Imerman
Imerman@iastate.edu
515.294.5781
Department of Economics
Iowa State University
Ames, Iowa 50011

I would like to thank Liesl Eathington from the Iowa State University Department of Economics for helpful comments on the content and exposition of this narrative.

Page updated April 24, 2006