

# Comparing Job Characteristics from the 2010 SIPP-EHC Field Test to the Census Bureau Business Register\*

Martha Stinson, Graton Gathright, and Jeremy Skog  
U.S. Census Bureau

March 20, 2012

## 1 Introduction

The Census Bureau collects information on households and businesses through mostly separate processes. With the redesign of the Survey of Income and Program Participation (SIPP), an effort has begun to evaluate household survey data on jobs by comparing them to corresponding reports made by employers. The purpose of this comparison is two-fold: first, to determine what improvements can be made to the survey questionnaire to more accurately capture the desired income and employer characteristics; second, to document differences between the household and business data and begin to consider the impact of these differences for the various uses of the SIPP.

In this paper, we present initial results from comparing data from an early (2010) field test of the new SIPP survey instrument, the SIPP Event History Calendar (SIPP-EHC), to data from IRS Form W-2 and to the Census Bureau Business Register (BR). In order to perform such a comparison, we first use Social Security Numbers (SSNs) to match SIPP respondents to all their W-2 records for employment in calendar year 2009, the reference period of the survey. Employers file W-2 forms annually with the Internal Revenue Service (IRS) for each employee and so, theoretically, each W-2 record corresponds to a job held by the respondent. In addition to SSNs, W-2 records contain Employer Identification Numbers (EINs) which link businesses to the BR, which contains the name, address, and other characteristics of employers. The survey also attempts to collect information about individual jobs held, including the name of the employer and the physical address. In order to compare specific firm characteristics reported in the SIPP and the BR, we follow our person-level match with a job-level match using the employer name and address. Thus for all SIPP respondents with SSNs and W-2s, we attempt to match each reported job in the SIPP to the employing establishment in the BR. After accomplishing this link, we are able to compare firm type (i.e. single- or multi-unit), firm size (i.e. number of employees), annual earnings, industry coding, and class of worker classification between the survey and the BR.

---

\*We would like to thank Shawn Klimek, Javier Miranda, Paul Hanczaryk, Gary Benedetto, Judy Eargle, Jason Fields, Becky Chenevert, Marina Vornovytsky, and Dave Hedengren for helpful comments and suggestions. We are grateful to Bill Winkler for the use of the BigMatch software for probabilistic linking. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau. All of the data used in this paper are confidential data. The U.S. Census Bureau supports external researchers' use of some of these data items through the Research Data Center network ([www.census.gov/ces](http://www.census.gov/ces)). All comparative statements in this report have undergone statistical testing, and, unless otherwise noted, all comparisons are statistically significant at the 5 percent significance level. For more information on the SIPP including sources of sampling and nonsampling error and statistical uncertainty, please see the SIPP main website <http://www.census.gov/sipp/> and the Source and Accuracy Statement for the 2008 SIPP panel at <http://www.census.gov/sipp/source.html>.

Our initial results show that there are a smaller number of jobs reported to the SIPP-EHC than to the IRS via the W-2 form. The W-2 count may not necessarily represent the same concept as the survey-reported job count, yet, the differences indicate the possibility that some jobs are not being reported in the survey. We find higher earnings on average in the SIPP-EHC than on the W-2 records. We believe that this result is mostly the case because of extreme outliers in the survey earnings. We find substantial disagreement about firm type and size between SIPP-EHC and BR, some of which might possibly be explained by differences in timing of the firm reports on number of employees and differences in how some government agencies such as school districts are classified. For industry, when we compare at the NAICS 2-digit level, the SIPP-EHC and the BR agree about 75% of the time.

Our paper proceeds as follows. We begin with a detailed description of data from 2010 SIPP-EHC, IRS Form W-2, and the BR and of the linking process used for combining these data. We then present results from comparing person level employment measures and from job-level employer characteristics. We conclude with thoughts on future work.

## 2 Data

In our analysis, we evaluate data on employment from the 2010 field test of SIPP-EHC and the 2008 panel of SIPP. We link sample persons in the surveys to data from IRS Form W-2. We also link survey-reported jobs in 2010 SIPP-EHC to data about employers from the Census Bureau's Business Register.

### 2.1 Data from 2010 SIPP-EHC

From January to March of 2010, the Census Bureau fielded a test of the SIPP-EHC instrument. This was the first computer-aided personal interview (CAPI) test and the first test to start with a fresh survey sample. The new instrument is scheduled to be used for the first time with a full national SIPP sample in 2014.

The sample for the 2010 SIPP-EHC included 7982 units. The sample was not designed to be nationally representative, but was restricted to units from the unit frame, self-representing primary sampling units (PSU), the high-poverty within-PSU stratum, in one of ten particular states<sup>1</sup>, and serviced by one of six particular field offices<sup>2</sup>.

The unit response rate for the field test was 82% and these responses yielded data on 14,738 sample persons. Of these sample persons, 5,686 reported at least one job during the reference year.

The instrument was designed to collect a one year history for up to five jobs including up to two employment spells per employer, but the instrument did not constrain job details to match across spells within the same employer, so, in some cases the second spell may actually correspond to a different employer. We treated second spells as different jobs unless the employer names matched identically. Theoretically, then, up to ten different employers could be reported for a given sample person in the 2010 SIPP-EHC instrument. A clerical review of the spells found that there were only 42 second spells that we erroneously treated as different jobs.

From the unedited data from this 2010 SIPP-EHC field test, we used or created job-level and person-level measures for our analysis. At the job level, we created a measure of annual earnings and we used the following reported employer characteristics: industry, number of employees at the place-of-work, number of employees firm-wide, whether the employer operated at more than one location, and whether the employer was a government entity, and if so what kind. At the person-level, we created a measure of annual earnings and a count of jobs held during the one-year reference period.

Our job-level annual earnings measure was created from data on the timing of employment spells and the

---

<sup>1</sup>2010 SIPP-EHC states were California, Connecticut, Illinois, Louisiana, Massachusetts, Maryland, Rhode Island, Texas, Wisconsin, and New York.

<sup>2</sup>2010 SIPP-EHC field offices were Boston, Chicago, Dallas, Los Angeles, Philadelphia, and New York.

dynamics of hours and pay rate for each job. In the data, start and ending months were reported for all job spells; where start day and ending day of jobs spells were missing we replaced them with the first and last days of the month, respectively. We dropped reported within-spell changes in pay rate or hours if there was insufficient information about when the change occurred. We also dropped spells and changes for which pay rate was reported as hourly but no information about hours worked was available.

In some cases, the reported pay rate seemed implausibly high. These were often hourly pay rates that would be more plausibly interpreted as a monthly salary. To address the possibility that questions regarding pay rate had been misunderstood or the answers mis-keyed, “soft checks” for extreme values in hourly and annual pay rates were included in the next field test, 2011 SIPP-EHC. For 2012 SIPP-EHC, additional soft checks for weekly, bi-weekly, semi-monthly, and monthly pay rates will be introduced. In our 2010 SIPP-EHC earnings measure, we dropped spells for which the reported pay rate was beyond the threshold for triggering these soft checks. Similarly, we ignored changes to pay rates where the new pay rate exceeded the soft-check threshold.

We calculated our person-level annual earnings measure by summing annual earnings across all jobs for each sample person. In our person-level counts of all jobs held annually, we have treated pairs of spells that were clerically determined to be for the same employer as single jobs.

From SIPP-EHC, we also used industry codes for each employer. Respondents’ descriptions of what work they do for each job they report were coded to a Census Industry Code (CIC) using a process that involves some automation and some hand coding. The process is the same one used for current production SIPP, the Current Population Survey, and the American Community Survey. A crosswalk between CIC and North American Industry Classification System (NAICS) is maintained so that CIC codes can be converted to NAICS codes. CIC codes are four digit where as NAICS codes can be up to six digits. Some CIC codes only correspond to a three digit NAICS code, meaning that for these industry groups, the Census coding is not at the same level of detail as is possible in NAICS. Sometimes a CIC code will correspond to a subset of larger digit NAICS codes within a smaller digit NAICS group. For example, CIC 2390 (rubber products manufacturing, except tires) corresponds to NAICS 32622 and 32629. Thus any survey job coded with CIC 2390 is given two NAICS codes.

## **2.2 Data from SIPP**

In our person-level analysis we also use data from 2008 SIPP in order to compare the extent of agreement between administrative records and survey data across SIPP and SIPP-EHC. Accordingly, we restrict our analysis of 2008 SIPP to the units that match the 2010 SIPP-EHC sample design.

We created person-level measures of annual earnings and of the number of employers in the year from edited 2008 SIPP data. For the annual earnings measure, we sum non-imputed monthly earnings from employment. In order to match our annual earnings measure from SIPP-EHC, we code missing earnings as zero in months for which the sample person was interviewed. However, we code monthly earnings as missing that are missing due to wave non-response or due to a sample person joining the sample after wave 2. In our annual earnings measure, we deal with these missing earnings by weighting up the observed person-months for an individual to represent all 12 months of the year.

We create an annual job count from SIPP using the edited employer numbers that are designed to link a sample person’s employers across waves.

## **2.3 Data from IRS Form W2**

From IRS Form W-2 data for Calendar Year 2009, we create job- and person-level annual earnings and person-level measures of annual number of employers. Workers in these data are uniquely identified by

a Census Bureau Protected Identity Key (PIK), a confidentiality-protected version of the Social Security Number. Employers in these data are identified by an Employer Identification Number (EIN).

We create our EIN-level measure of annual earnings from these data by summing the value of the “Wages, Salary, and Tips” field with the value of the “Deferred Compensation” field. The person-level version was created by summing the EIN-level measure across EINs for each PIK. We count the number of distinct EINs for each PIK to create our person-level measure of annual number of jobs.

## 2.4 Data from the Census Bureau Business Register

We link job reports in 2010 SIPP-EHC to data about the employing establishments from the Census Bureau Business Register (BR). Establishments in the BR are operating units of an enterprise that typically correspond to a particular location of business. The BR contains data on these establishments which comes from the tax filings of these businesses and from their responses to various business surveys.

Some enterprises in the BR, called single-unit enterprises, operate only one establishment. Enterprises that operate more than one establishment are called multi-unit enterprises. Multi-unit enterprises often use more than one EIN, the employer identifier on each W2 record, to report economic activity to the IRS about sub-groups of establishments.

The BR records a (potentially two-part) name and two addresses (physical and mailing) for each establishment. A (potentially two-part) name is also sometimes recorded at the EIN level and is called the “submaster” name.

Each establishment in the BR, then, is classified as belonging to a single- or multi-unit enterprise. The BR also records a five digit NAICS code describing the industry of each establishment.

For each multi-unit enterprise, the BR records the number of employees at each establishment and in the entire enterprise as of March 12 of each year, derived from business survey data and IRS tax reports. We use these counts as our measures of size for multi-unit establishments and enterprises. For single-unit enterprises, the number of employees at the enterprise/establishment is recorded in the BR based on quarterly tax filings. We use these data as our measure of firm size for single-unit enterprises.

Enterprises in the BR are also classified as government or non-government.

## 2.5 Data Linking

We perform two types of data-linking for our analysis. First, we link respondents in 2010 SIPP-EHC and in 2008 SIPP to IRS Form W-2 data using a unique person identifier that was added to both datasets using standard Census Bureau processes that are documented elsewhere. This identifier is a protected version of the Social Security Number and is called the Protected Identity Key (PIK).

For some sample persons there is no PIK available. The PIK rates in 2010 SIPP-EHC and the geo-matched 2008 SIPP for sample persons aged at least 15 years at time-of-interview are 81.9% and 74.4%, respectively. In 2010 SIPP-EHC, there were 5,686 sample persons reporting at least one job. Of these, 4,379 had a PIK available for linking and 3,740 were linked by PIK to at least one W-2 record.

There were 6,540 jobs reported in 2010 SIPP-EHC, with 5,083 jobs held by sample persons for which a PIK was available. For 4,377 of these, there was at least one potentially matching W-2 record.

Matching reported jobs to specific W-2’s and to the employing establishments was the second type of linking that we performed. Using probabilistic matching on name and address of employer, we linked SIPP-EHC reported jobs to establishments in the BR. The establishments to which a given job could be matched were all of the establishments belonging to an EIN that had issued a W-2 for the sample person reporting the job.

We performed the probabilistic matching using BigMatch, software developed at the U.S. Census Bureau. We employed multiple passes using different linking variables, parameter values, and score cutoffs. The parameter values and score cutoffs were chosen by clerical review of sets of matched and non-matched pairs. The passes were performed in descending order of stringency of criteria. After excluding any matching results that did not meet pass-specific score cutoffs, we chose, as the linked establishment for each job, the paired establishment with the highest match score in the earliest pass. Prior to the probabilistic matching passes, we cleaned and standardized the data using the SAS Data Quality Server and additional custom routines. documentation on what this processing does to the data.

There were 468 jobs that did not appear in any pair with a score sufficiently high to exceed the cutoff in any pass. For these, we clerically reviewed the possible matches; by this method we were able to match an additional 79 jobs to establishments.

### 3 Results

We present comparisons of information about jobs between the 2010 SIPP-EHC data and administrative records, first at the person-level and then at the job-level. We first compare fact-of-employment between the survey and the W-2 records. Also at the person-level, we compare counts of the number of jobs held and annual earnings. To benchmark the performance of the SIPP-EHC, we perform these comparisons for the geo-matched 2008 SIPP as well. At the job-level, we compare survey reports with BR records on firm organization, public versus private employment, establishment and enterprise size, and industry. In many instances, we are cautious about declaring the administrative data to be truth, therefore we simply highlight the differences between the two data sources.

#### 3.1 Employment, Earnings, Job Counts

##### 3.1.1 SIPP-EHC and W-2 Comparison

We begin our comparison of survey job reports and W-2 data by looking at the percentage of people who have job records in both sources. In Table 1, we group individuals by whether they reported a job in the SIPP-EHC and whether they had any W-2 records for 2009, the survey reference year. Of those with no reported jobs in the SIPP-EHC, 13% had at least one W-2 record. Of those who reported a job in the SIPP-EHC, 85% had at least one W-2 record, while approximately 15% had no W-2 records. This last group represents at least two types of jobs: self-employment and informal jobs where payments were not reported to the IRS by the employer. Next, considering the universe of individuals without W-2 records, it appears that about 10% have some type of earnings. For individuals with at least one W-2 record, 81% report some job activity in the survey. These results lead us to suspect that the SIPP-EHC is failing to capture information about a non-trivial number of jobs. However this table also highlights the value of survey data collection for insights into jobs that are not reflected in tax records.

In Table 2, we present average earnings for each cell in Table 1 in order to begin to paint a picture of the types of jobs we are missing. For those with no job report in the SIPP-EHC, we report average W-2 earnings when there are W-2 reports. This amount (\$13,684) is the total amount earned and reported on all W-2 records for an individual in 2009, averaged across individuals in the SIPP-EHC and reported in the top right cell of Table 2<sup>3</sup>. This amount leads us to believe that we are missing some jobs with significant enough earnings that total earnings for these households will be artificially low. For individuals with SIPP-EHC jobs but no W-2 records, we report average survey earnings as \$36,127. This amount is the total amount earned and reported in the SIPP-EHC for an individual in 2009, again averaged across all individuals. Here we conclude

---

<sup>3</sup>We remind the reader that our sample is not nationally representative and hence these average amounts will not compare to published national averages. We have not adjusted these amounts for inflation. All amounts are in 2009 dollars, as they were reported.

that jobs which do not produce W-2 forms nonetheless report positive amounts of income. Capturing this income is one of the strengths of the survey relative to the administrative records.

In the bottom-right cell of Table 2, we show both average W-2 and survey earnings for individuals who had jobs reported in both sources. Contrary to our expectations from past studies of SIPP survey earnings, we found that SIPP-EHC earnings were higher on average than W-2 earnings for individuals with job reports in both the survey and the tax forms. The difference of approximately \$8000 was significant at the .1% level<sup>4</sup>. Investigation into the cause of this has led us to conclude that some respondents gave unreasonable answers when asked about their rate of pay. In particular, respondents who said they were paid by the hour sometimes reported hourly wages greatly in excess of an amount that would produce a reasonable amount of annual earnings. In spite of our attempt to discard the most egregious reports, enough outliers still remain to skew average earnings.

In order to further investigate our hypothesis that outliers are producing this result, we present median earnings for these same three cells in Table 3. The median is lower than the mean for each cell for both types of earnings, with skewness coefficients of 5.3 for both types of SIPP-EHC earners and 4.1 and 2.5 respectively for W-2 earners with and without survey earnings. Thus we conclude that survey non-reporters are a group with some high earners who raise the average W-2 earnings to about twice as much as the median. For those who do report earnings in the survey, median SIPP-EHC earnings are approximately \$14,000 lower than average SIPP-EHC earnings for those with W-2 reports and \$20,000 lower for those without W-2 reports. This lends credence to our theory that outliers are the main cause of disagreement between average W-2 earnings and average SIPP-EHC earnings. Median SIPP-EHC earnings are only approximately \$2,000 higher than median W-2 earnings for sample persons with jobs reported in both sources. Data from the 2011 and 2012 SIPP-EHDC field tests will permit an evaluation of the improvement to earnings measurement from implemented instrument changes such as real-time checks on the level of monthly earnings implied by answers to pay rate and hours worked questions.

In Table 4, we compare the number of jobs reported in the SIPP-EHC to the number of jobs producing W-2 records for those with job reports in both sources. Before examining these percentages, it is worth considering what the W-2 job count actually represents. Employers file W-2s with the IRS using employer identification numbers (EINs) and it is possible for a multi-unit company to have multiple EINs and to classify different parts of the total amount earned by a given employee under different EINs. It is also possible for a company to change ownership during a year and acquire a new EIN, which would then produce a second W-2 for every employee. We have done some preliminary work on both these fronts to try to determine the extent to which such arrangements are present in our data and we do not believe that these types of occurrences are widespread. However, this work is still preliminary and we have not been able, at this point, to produce a W-2 job count that combines records which a survey respondent would most likely view as coming from one employer. Hence the percentages in this table should be viewed as noisy measures. Nonetheless, there are some general patterns of interest. Of individuals who report only one job in the SIPP-EHC, approximately 20% have two or more W-2 records (sum of row percentages for categories above one job). Of individuals with two W-2 records, 65% (column percentage) report only one job in the survey and more than half of individuals with three, four, and five or more W-2 records report only one job in the survey. As in Table 1, we conclude that it is likely that SIPP-EHC respondents are failing to report some jobs.

### 3.1.2 2008 SIPP, SIPP-EHC, and W-2 Comparison

In order to place these comparisons between W-2 and SIPP-EHC earnings and job counts into context, we also compare the geo-matched<sup>5</sup> 2008 SIPP panel to W-2 records. Unlike the SIPP-EHC, in the 2008 panel, employer jobs and self-employment were collected in separate parts of the instrument. Hence for this comparison between the two surveys, we restrict attention to jobs reported as being for an employer,

---

<sup>4</sup>The t-stat for this comparison was -8.5294 with degrees of freedom equal to 2684.

<sup>5</sup>See Section 2.

as “other arrangement”, or as “contingent work.” We begin by reproducing Table 1, but not counting self-employment jobs. As shown in Table 5, we see that this has almost no effect on the percentage of people with no reported employment in the SIPP-EHC who nevertheless match to W-2 records. However the percentage with SIPP-EHC employment and no W-2 records drops by five percentage points, from 14.5% to 9.3%. Self-employment appears to explain about a third of the cases where a person reported working but had no W-2. In Table 6, we show the same results for the 2008 SIPP and see that the overall picture of survey employment and W-2 reports is comparable between the two survey sources. Approximately 10% of respondents with reported jobs have no W-2 records while 11% have no survey employment but do have W-2 records. This later group is about three percentage points smaller than in the SIPP-EHC.

In Table 7, we present comparisons of differences between survey and W-2 measures for annual earnings and for annual job counts across 2008 geo-matched SIPP and 2010 SIPP-EHC. We regress differences between the administrative and survey versions of each measure on a constant and an indicator for inclusion in the SIPP-EHC (SIPPEHC=1) or SIPP (SIPPEHC=0) survey. In the first column, we calculate the difference between total 2009 survey earnings and total 2009 W-2 earnings at the person level and then for all individuals with positive W-2 earnings. The constant represents the average difference between tax records and survey reports for the geo-matched 2008 SIPP sample is approximately -\$4000. Since the coefficient is negative, the W-2 earnings are higher and the difference is significantly different from zero. The coefficient on the SIPP-EHC indicator is not significantly different from zero and we conclude that for this broad group, there is a significant difference between survey and tax earnings but this difference does not vary across the two surveys in a statistically significant way.

In the second column of Table 7, we restrict ourselves to comparing only respondents with positive earnings in both survey and tax records. The average difference drops for 2008 SIPP respondents to -\$2900. Given that we dropped individuals who had zero earnings in the survey, this change is not surprising. The difference between zero earnings and positive earnings is larger on average than the difference between two positive earnings reports. However, the coefficient on the SIPP-EHC indicator is now positive and quite large, approximately \$11000. This reflects our earlier findings that average SIPP-EHC earnings were substantially higher than average W-2 earnings, a feature not present in the SIPP 2008 data.

In the third column of Table 7, we look at average differences in job counts. Relative to their W-2 job counts, respondents in the 2008 SIPP report approximately a third of a job less. Respondents in the SIPP-EHC report almost a half a job less (sum of constant and SIPP-EHC indicator coefficients). The coefficient on the SIPP-EHC indicator is statistically significant, meaning that this difference between job counts in the 2008 SIPP and the SIPP-EHC is significantly different from zero. Since the 2008 SIPP panel interviewed respondents every four months, the job count is dependent on the correct tracking of jobs across waves. If job reports from multiple interviews are not properly connected with a common employer identifier, the job count will be artificially inflated. Historically there have been many difficulties in linking employer reports across waves in the SIPP and while many problems have been resolved, we believe that some still remain. Hence it is unclear if the SIPP-EHC is really worse than the 2008 SIPP in collecting job counts or whether there are simply fewer seam issues. In any case, both surveys collect fewer jobs than are found in the W-2 records.

### 3.2 Job Characteristics

We now turn to job-level comparisons between the survey and administrative data. Because the SIPP-EHC collected the identifying information for employing establishment including place-of-work address, we were able to compare establishment level job characteristics in a way not heretofore possible for SIPP<sup>6</sup>. These comparisons take advantage of the link between survey-reported jobs and specific establishments in the BR that is discussed in Section 2. As reported in Table 8, we were able to match 69.6% of the 4,377

---

<sup>6</sup>The American Community Survey collects both name and address of the employer but no SIPP panel prior to the SIPP-EHC has collected this information.

survey-reported jobs for sample persons having a PIK and at least one W-2 record.

We had not expected to match jobs that were reported as self-employment for an incorporated business paying a salary to the worker. However, of these 105 unlikely-to-match jobs, 22 did match to an establishment in the BR. This suggests the possibility of some mis-reporting about the type of job. For jobs that were expected to match to the BR, the match rate was 70.8%. Jobs may have failed to match because of insufficient or poor quality information for identifying the employing establishment, or because the employment did not generate a W-2 record. Some reported jobs have no W-2 record because the employer failed to report the employment to the IRS. Other reported jobs have no W-2 record because the reported job was actually the type of self-employment that generates no W-2 record.

### **3.2.1 Company Organization: Number of Locations and Private versus Government Sector**

Both the survey and BR data furnish measures of whether an employer operates at only one or at more than one location (see Section 2). In Table 9, we present a job-level comparison between the survey report of whether the employer operates at more than one location and the BR classification of the enterprise as single-unit or multi-unit. Of the jobs reported to be for employers operating in multiple location, 62.51% matched to establishments of multi-unit enterprises; the remaining 37.49% matched to single-unit enterprises. Agreement between survey and administrative data was stronger for jobs reported to be for single-location employers: 75.54% of these jobs matched to single-unit enterprises.

The discrepancies between survey-reported and BR-recorded number of employer locations may indicate mis-reporting by survey respondents. Some of the disagreement may also arise from differences between the concepts of location in the two measures. For instance, it appears that in public sector employment, business entities likely to be operating in multiple locations are sometimes recorded in the business register as single-unit enterprises. In particular, school districts may operate multiple schools but be recorded as single-unit enterprises in the administrative data.

In Table 10, we present a comparison between the survey and BR on whether the employer was in the private or government sectors. Of jobs reported in the survey as being for the federal government, 72.5% matched to a government enterprise in the BR. Some of this disagreement about employment for the federal government may arise from mis-reporting about type of employer by government contractors responding to the survey. Agreement between the survey and BR about the employer being a government entity was strong for jobs reported as state government (86.4%), local government (90.5%) and armed forces (87.5%).<sup>7</sup> Nearly all (98.6%) of the jobs reported to be for for-profit private-sector employers matched to private-sector enterprises in the BR. Agreement between survey and BR about government status was 88.7% for jobs reported to be in the non-profit private sector.

### **3.2.2 Establishment and Enterprise Size**

We compare survey-reported information about the number of employees at employing establishments and enterprises<sup>8</sup> with counts of employees at establishments and enterprises in the BR. As described in Section 2, the survey measure of establishment and enterprise size (number of employees) is collected as one of 5 mutually exclusive collectively exhaustive size ranges in which the employee count would fall. As a convenience in comparing these data to the BR, we similarly discretize the employee counts recorded in the BR.

For different types of enterprises in the BR, we use different BR measures of establishment and enterprise size. For single-unit establishments/enterprises, we use the quarterly employment measures in the BR that come from quarterly tax filings and represent number of employees present at the end of a tax quarter. In our comparisons, we use the size measures for the last quarter when the respondent was employed by

---

<sup>7</sup>Only one job reported as being for the armed forces was matched to a non-government enterprise in the BR.

<sup>8</sup>We use the term enterprise to mean the employing enterprise.

the enterprise. For multi-unit enterprises and the establishments of multi-unit enterprises, we use the BR employee counts from business surveys that are intended to measure establishment and enterprise sizes on March 12th.

In Table 11, we present a comparison of the survey and BR establishment size measures for jobs identified in the BR as single-unit enterprises. Of jobs matched to single-unit enterprises with less than 10 employees according to the BR, 76.0% reported in the survey in the same size range. Another 16.7% were reported to have between 10 and 24 employees, inclusive. Some of the jobs for which the survey and BR measures differ may reflect the differences in the count dates across the measures rather than respondent mis-reporting.

For jobs matched to single-unit enterprises with 10 to 24 employees, the percentage with the survey report of enterprise size in the same range as the BR measure was 52.7%. Another 28.6% of these jobs fell in the next lower size range and 15.8% in the next higher size range. A very similar pattern was found for the jobs matched to single-unit enterprises with 25 to 99 employees. For the jobs matched to single-unit establishments in the 100 to 499 range, the jobs were spread more evenly across the matching survey reported size range and then one below. The two highest BR-recorded size ranges have the lowest level of correspondence to the survey reported size range. For firms in the 1000+ category in the BR, only 17.8% are reported to be in that category in the SIPP-EHC. The largest number of 1000+ employee firms are reported to be in the 25-99 and 100-499 categories. The same pattern is true for the 500-999 BR size category which agrees with the SIPP-EHC report in only 12.5% of cases and has approximately 67% of survey reports in the two categories below the BR size.

Across all of the BR enterprise size categories, substantial numbers of jobs which matched to BR single-unit enterprises reported a lower enterprise size category than recorded in the BR. In future work, we would like to investigate to what extent these were jobs for which the company organization (number of locations) was reported as multi-unit. If the respondent viewed her place of work as a multi-unit enterprise, then it is not unreasonable that her establishment size report would only account for part of the enterprise's total employment. As mentioned above, we found that school districts likely to be operating at multiple locations were often recorded in the BR as single-unit establishments.

In Table 12, we begin to look at size comparisons for multi-unit firms, beginning with a cross-tabulation of survey and BR measures of establishment size (i.e. work location). Of jobs matched to multi-unit enterprises with less than 10 employees at the establishment according to the BR, only 39.3% have survey-reported size in that same size range. Another 22.1% were reported to have between 25 and 99 employees, inclusive. Nearly 30% is split evenly between the 10 to 24 and 100-499 categories. Perhaps the partitioning of enterprises into units for reporting purposes is finer than respondents perceive based on the notion of "location." Some of this disagreement may also reflect the differences in the dates to which the counts correspond across rather than mis-reporting by respondents.

For jobs matched to establishments belonging to multi-unit enterprises and having establishment-level employment in the 10 to 24 employees range, 35.6% had a survey report in that same establishment size range. Another 28.8% were reported to be in the 25 to 99 employee range. For the jobs matching to establishments of multi-unit enterprises having establishment size in the ranges 25 to 99 and 100 to 499, approximately 51% of matching jobs had survey reported establishment size that agreed with the BR counts, and approximately 20% reported establishment size in the next lower category.

Of the jobs matched to establishments belonging to multi-unit enterprises and having establishment-level employment in the 500-999 range, approximately 60% are roughly split between the survey-reported establishment size ranges of 100 to 499 and 500 to 999. Another 24.7% were reported in the survey to have establishment-level employment in excess of 1000 employees. Nearly 50% of the jobs matched to establishments belonging to multi-unit enterprises and having establishment-level employment in excess of 1000 employees had survey-reported establishment size that agreed. Another approximately 30% was roughly split between the next two lower size ranges. Greater agreement for these size ranges may reflect that intra-year fluctuation in establishment size may be less likely to move an establishment out of these larger size ranges.

It may also be easier for survey respondents to correctly place employers in these large ranges.

In Table 13, we move to size comparisons at the enterprise level which utilizes the survey question asking about the number of employees at all locations of the employer. We present a cross-tabulation of enterprise-level employment as reported in the survey and recorded in the BR using the BR employment counts as of March 12. Of the survey-reported jobs that match to multi-unit enterprises, 75.1% match to enterprises with over 1000 employees, 82.3% of which reported as much in the survey. For the jobs matched to multi-unit enterprises in next three lower enterprise size ranges, there appears to be substantial over-estimation of enterprise-wide employment by survey respondents. Nearly half of the jobs matched to enterprises in the 500 to 999 range were reported as being for enterprises with 1000 or more employees. For the 25 to 99 and 100 to 499 ranges, the corresponding percentages were 24.2 and 27.7, respectively. Only two multi-unit enterprises had less than 25 employees, both of these had more than 10 employees.

### 3.2.3 Industry

We next compare the survey-reported industry of the employing establishments with the industry of establishments as recorded in the BR. We perform multiple comparisons at increasing level of detailed coding, specifically, at NAICS digit-levels 2 to 5. Also, as discussed in Section 2 above, survey-reported industry is sometimes coded as corresponding to a set of higher-digit NAICS codes within a smaller-digit NAICS code. In our comparisons, we consider the survey-reported industry to agree with the BR if any of the NAICS codes corresponding to the survey-reported industry matches the NAICS code recorded in the BR for the establishment to which the job is matched.

In Table 14 we present rates of agreement between survey report and BR on industry coding at digit levels 2 to 5 for jobs that matched to single-unit enterprises. At the least-detailed, two-digit level, survey-response-based coding agreed with the BR for 74.1% of the jobs. Each successive higher-digit-level comparison excludes more jobs since survey-reported industry is coded only to the detail level that the CIC to NAICS crosswalk allows. The agreement rates at the 3-, 4- and 5- digit levels are 67.0%, 62.9%, and 60.4%, respectively.

In Table 15, we present the corresponding results for jobs that matched to establishments of multi-unit enterprises. At the two-digit level, the survey-report-based NAICS coding agreed with the BR for 72.70% of these jobs. The agreement rates at the three-, four- and five-digit levels are 66.6%, 56.0%, and 47.3%, respectively. Since industry classification can differ across establishments within an enterprise, it is possible that some of the agreement about industry classification for jobs matched to establishments of multi-unit enterprises is due to matching jobs to the incorrect establishment with an enterprise.

It is also important to note that when asked about what an employing firm does, the respondent is asked to classify the firm into a broad category (manufacturing, wholesale trade, retail trade, service, or something else) and then to give the main activity of the firm. If the employer is a multi-unit enterprise, then the respondent is asked to report the main activity of the enterprise at the location where the sample person is employed by the enterprise. If the firm is engaged in multiple activities at the work location, then the respondent is asked to choose the main one. Since people are best informed about what they personally do, some respondents may have trouble reporting a main activity and instead simply report their own activity. This would cause disagreement between the survey job report and the BR industry code in a way that is different from standard mis-reporting models. For instance an individual who is employed by a restaurant but works in the gift shop may be classified by the Census industry coding process as belonging to the retail sector when, in fact, his employer is classified as belonging to the restaurant service industry. We may be hesitant to declare the worker's response to be "wrong" in this case. However, the output from this employer, i.e. the product of the worker, is classified in the national product accounts under the service industry. More research is needed to determine the effects on analyses from classifying workers according to their firms' industry designations or their own.

In Table 16, we present the distribution across two-digit BR NAICS code of jobs matched to single-unit

enterprises for which the industry coding does not agree between survey-report and BR. The highest rates of disagreement were in “Wholesale Trade” (12.57%), “Admin. Support, Waste Management” (13.90%), and “Public Admin” (16.84%). Wholesale trade is probably difficult for survey respondents to describe in a way that does not sound like retail trade or manufacturing. The Admin. Support and Waste Management category includes temporary employment agencies for which some respondents may be describing the work that they perform at their job in the firm where they are placed, but this would not match to the temporary employment industry coding. Public Administration had the highest rate of disagreement suggesting that the jobs are hard to code correctly or the industry covers many different functions - schools are sometime coded as Public Administration instead of education in the BR, for example.

Table 17 presents the corresponding distribution for jobs matched to establishments of multi-unit enterprises. The highest rates of disagreement were in “Management of Companies” (14.29%), “Admin. Support, Waste Management” (12.24%), and “Health Care, Social Assistance” (10.46%). It may be that employees working at a company’s headquarters but whose activity is something quite different from management are generating the high error rate in the Management of Companies category. For both single and multi-unit firms, it appears that industry disagreement is not entirely random and there are some categories that are either harder for respondents to describe or harder for Census to code.

Finally, we present the distribution of jobs over two-digit level survey-report-based industry classifications for jobs for which we found no matching establishment in the BR. We split this group into two categories: 1) jobs for sample persons with no W-2 record and 2) jobs for sample persons with at least one W-2 record. Results for the first group are presented in Table 18. Clusters of non-matching jobs for persons with no W-2s are in “Construction” (15.13%), “Health Care, Social Assistance” (10.96%), and “Other Service” (12.17%) which includes dry cleaning, beauty salons, and car washes. Other industry categories with a substantial number of these non-matching jobs are “Accommodations, Food Service” (7.48%), which includes hotels, and “Admin. Support, Waste Management” (8.69%), which includes landscaping companies. These are industry categories where it would not be surprising to discover employment that is not generating W-2 records. Collecting information on such employment is one of the strengths of household survey data.

Table 19 presents similar results for unmatched jobs for sample persons who had a least one W-2 record. The industry category with the largest clustering of such jobs is “Health Care, Social Assistance” (15.23%). Other industries with clusters are “Educational Services” (9.36%) and “Admin. Support, Waste Management” (7.80%). Some of these non-matches arise from jobs that are not closely tied to a specific location and are therefore more difficult to match to a specific establishment. However others are likely from jobs that are not generating W-2 records. We conclude that the survey provides a useful picture of the informal labor market and captures some earnings that would be missed by using administrative data alone.

## 4 Conclusion

We have documented instances where the SIPP-EHC reports on jobs and W-2 records disagree. In general, W-2 records produce higher job counts and lower earnings. Reports about firm organization often does not agree with the BR for those reporting work at multi-unit firms. Firm size is difficult to measure and to report in the small size categories, with many reports being just above or below the BR category. In larger sizes for multi-units, there is more agreement while in larger sizes for single-units there is less agreement and less of a clear pattern. When comparing industry, we find that three quarters of the jobs have the same two digit NAICS from the BR and the SIPP-EHC but this agreement decreases as we compare at greater levels of detail. Many of the industry disagreements are concentrated in Wholesale trade, Public Administration, Firm Management, and the sector which includes temporary help firms.

Not all of this disagreement is necessarily the result of reporting error by SIPP-EHC respondents. Work remains to be done to investigate the structure of firms in the BR (especially in the government sector) and to create a more accurate job count from the W-2 records that reflects parent ownership and succes-

sor/predecessor relationships. There is also the need for more detailed industry comparisons and investigation into how differences in household and firm industry reports affect analyses using industry classifications.

In spite of these qualifications on the need for more careful investigation on data issues, we feel that this work sheds light on areas of further improvement for the SIPP-EHC. We are hopeful that the 2011 SIPP-EHC field test will show substantial improvement in the reported earnings. We also feel that it is likely that the survey is missing jobs and so we continue to consider ways to collect a more complete picture of each respondent's employment situation. We anticipate performing this evaluation for both the 2011 and 2012 SIPP-EHC field tests and continuing to learn about differences between the survey and the Business Register and the causes of these differences.

Table 1: Percentage of SIPP-EHC Respondents with PIKs reporting jobs versus linking to W-2 Records

Jobs Reported in SIPP-EHC	Matched to W-2 Records		
	No row%/col%	Yes row%/col%	Total row%/col%
No	86.64	13.36	100.00
	89.76	18.77	59.62
Yes	14.59	85.41	100.00
	10.24	81.23	40.38
Total	57.55	42.45	100.00
	100.00	100.00	100.00
<i>N</i>	10845		

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 2: Average Annual Survey and W-2 Earnings for SIPP-EHC Respondents with PIKs by reporting and linking status

	No_W2_Jobs W2_Earn/SIPPEHC_EARN	W2_Jobs W2_Earn/SIPPEHC_EARN
No_SIPPEHC_Jobs	.	13684.37
SIPPEHC_Jobs	36127.85	29515.53 38967.43
Observations	324	4604

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 3: Median Annual Survey and W-2 Earnings for SIPP-EHC Respondents with PIKs by reporting and linking status

	No_W2_Jobs	W2_Jobs
	W2_Earn/SIPPEHC_EARN	W2_Earn/SIPPEHC_EARN
No_SIPPEHC_Jobs		5934.00
		.
SIPPEHC_Jobs	.	22721.50
	15642.86	24611.43
Observations	324	4604

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 4: SIPP-EHC Job Counts by W-2 Job Counts

SIPP-EHC Job Count	W-2 Job Count						Total row%/col%
	1 job row%/col%	2 jobs row%/col%	3 jobs row%/col%	4 jobs row%/col%	5+ jobs row%/col%		
1 job	80.13	14.48	3.45	1.19	0.75	100.00	85.32
2 jobs	33.12	48.10	13.50	3.80	1.48	100.00	12.67
3+ jobs	24.00	25.33	30.67	10.67	9.33	100.00	2.01
Total	73.05	18.96	5.27	1.71	1.02	100.00	100.00
<i>N</i>	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	3740						

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 5: SIPP-EHC Respondents with PIKs reporting employer and contingent jobs versus linking to W-2 Records

Emp./Cont. Work Jobs Reported in SIPP-EHC	Matched to W-2 Records		
	No row%/col%	Yes row%/col%	Total row%/col%
No	86.10	13.90	100.00
	93.99	20.57	62.82
Yes	9.30	90.70	100.00
	6.01	79.43	37.18
Total	57.55	42.45	100.00
	100.00	100.00	100.00
<i>N</i>	10845		

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 6: SIPP-2008 Respondents with PIKs reporting employer and contingent jobs versus linking to W-2 Records

Emp./Cont. Work Jobs Reported in SIPP	Matched to W-2 Records		
	No row%/col%	Yes row%/col%	Total row%/col%
No	88.93	11.07	100.00
	92.63	14.83	58.59
Yes	10.02	89.98	100.00
	7.37	85.17	41.41
Total	56.25	43.75	100.00
	100.00	100.00	100.00
<i>N</i>	6291		

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation 2008 Panel

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 7: Comparisons of Average Annual Earnings and Job Counts between SIPP-2008, SIPP-EHC, and W-2s

	earn1	earn2	jobs
sippehc	-152.1 (759.4)	11153.3*** (1060.4)	-0.168*** (0.0246)
_cons	-4020.0*** (382.1)	-2873.5*** (426.7)	-0.302*** (0.0203)
<i>N</i>	7353	4954	7298

Standard errors in parentheses

Sources:

U.S. Census Bureau:

SIPP 2008 Panel and Event History Calendar Field Test 2010

Internal Revenue Service:

Form W-2 Records Tax Year 2009

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table 8: Job Match rate between SIPP-EHC and W-2

W-2 Match Expectations	Match Success		
	No match rcpt/cpct	Yes match rcpt/cpct	Total rcpt/cpct
Not expected to match to W-2	79.05	20.95	100.00
	6.24	0.72	2.40
Expected to match to W-2	29.21	70.79	100.00
	93.76	99.28	97.60
Total	30.41	69.59	100.00
	100.00	100.00	100.00
<i>N</i>	4377		

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 9: Firm Organization Comparison: Single versus Multi Unit

SIPP-EHC: more than one firm location	Firm Organization type in Business Register		
	Multi-unit row%/col%	Single-unit row%/col%	Total row%/col%
Yes	62.51	37.49	100.00
	81.96	46.88	64.00
No	24.46	75.54	100.00
	18.04	53.12	36.00
Total	48.81	51.19	100.00
	100.00	100.00	100.00
<i>N</i>	2817		

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Business Register

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 10: Government Status Comparison

SIPP-EHC Class of Worker	BR Govt. Indicator		
	Non-government row%/col%	Government row%/col%	Total row%/col%
Federal Govt	27.50	72.50	100.00
	0.44	7.77	1.39
State Govt	13.56	86.44	100.00
	0.64	27.35	4.11
Local Govt	9.52	90.48	100.00
	0.72	45.84	6.58
Armed Forces	12.50	87.50	100.00
	0.04	1.88	0.28
Unknown Govt type	71.43	28.57	100.00
	0.20	0.54	0.24
Private, for profit	98.61	1.39	100.00
	88.23	8.31	77.85
Private, non-profit	88.69	11.31	100.00
	9.73	8.31	9.54
Total	87.01	12.99	100.00
	100.00	100.00	100.00
<i>N</i>	2871		

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Business Register

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 11: Single Unit Firm Size Comparison: SIPP-EHC employment at location worked versus BR employment at end of last quarter worked at firm

SIPP-EHC	Business Register						Total
	1-9	10-24	25-99	100-499	500-999	1000+	
	row%/col%	row%/col%	row%/col%	row%/col%	row%/col%	row%/col%	row%/col%
1-9	51.69	17.85	10.15	7.69	1.85	10.77	100.00
	76.02	28.57	11.34	10.50	12.50	10.74	24.49
10-24	14.07	40.68	28.14	7.60	0.76	8.75	100.00
	16.74	52.71	25.43	8.40	4.17	7.06	19.82
25-99	2.99	8.70	39.67	19.02	5.16	24.46	100.00
	4.98	15.76	50.17	29.41	39.58	27.61	27.73
100-499	1.31	1.75	13.10	39.30	5.68	38.86	100.00
	1.36	1.97	10.31	37.82	27.08	27.30	17.26
500-999	2.99	2.99	4.48	34.33	8.96	46.27	100.00
	0.90	0.99	1.03	9.66	12.50	9.51	5.05
1000+	0.00	0.00	6.67	13.33	2.67	77.33	100.00
	0.00	0.00	1.72	4.20	4.17	17.79	5.65
Total	16.65	15.30	21.93	17.94	3.62	24.57	100.00
	100.00	100.00	100.00	100.00	100.00	100.00	100.00
<i>N</i>	1327						

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Business Register

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 12: Multi Unit Firm Size Comparison: SIPP-EHC employment at location worked versus BR employment at establishment on March 12

SIPP-EHC	Business Register						Total
	1-9	10-24	25-99	100-499	500-999	1000+	
	row%/col%	row%/col%	row%/col%	row%/col%	row%/col%	row%/col%	row%/col%
1-9	47.83	15.65	13.04	15.65	4.35	3.48	100.00
	39.29	11.25	4.62	4.62	6.17	2.52	9.16
10-24	11.24	32.02	38.20	11.24	0.56	6.74	100.00
	14.29	35.62	20.92	5.13	1.23	7.55	14.18
25-99	9.28	13.77	49.70	22.75	1.80	2.69	100.00
	22.14	28.75	51.08	19.49	7.41	5.66	26.61
100-499	5.73	5.73	15.47	57.31	7.16	8.60	100.00
	14.29	12.50	16.62	51.28	30.86	18.87	27.81
500-999	6.80	1.94	4.85	35.92	23.30	27.18	100.00
	5.00	1.25	1.54	9.49	29.63	17.61	8.21
1000+	3.98	9.66	9.66	22.16	11.36	43.18	100.00
	5.00	10.62	5.23	10.00	24.69	47.80	14.02
Total	11.16	12.75	25.90	31.08	6.45	12.67	100.00
	100.00	100.00	100.00	100.00	100.00	100.00	100.00
<i>N</i>	1255						

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Business Register

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 13: Multi Unit Firm Size Comparison: SIPP-EHC employment at all locations versus BR employment at enterprise on March 12

SIPP-EHC	Business Register					Total
	10-24	25-99	100-499	500-999	1000+	
	row%/col%	row%/col%	row%/col%	row%/col%	row%/col%	row%/col%
1-9	0.00	9.09	27.27	9.09	54.55	100.00
	0.00	3.03	2.19	1.64	0.85	1.18
10-24	5.88	11.76	17.65	5.88	58.82	100.00
	50.00	6.06	2.19	1.64	1.42	1.82
25-99	1.61	20.97	25.81	11.29	40.32	100.00
	50.00	39.39	11.68	11.48	3.56	6.63
100-499	0.00	6.77	51.13	7.52	34.59	100.00
	0.00	27.27	49.64	16.39	6.55	14.22
500-999	0.00	0.00	15.52	20.69	63.79	100.00
	0.00	0.00	6.57	19.67	5.27	6.20
1000+	0.00	1.22	5.81	4.59	88.38	100.00
	0.00	24.24	27.74	49.18	82.34	69.95
Total	0.21	3.53	14.65	6.52	75.08	100.00
	100.00	100.00	100.00	100.00	100.00	100.00
<i>N</i>	935					

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Business Register

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 14: Industry Comparison: Single-unit

	NAICS agree 2 digit	NAICS agree 3 digit	NAICS agree 4 digit	NAICS agree 5 digit
	pct/obs	pct/obs	pct/obs	pct/obs
No	25.89	32.98	37.06	39.64
Yes	74.11	67.02	62.94	60.36
Total	100.00	100.00	100.00	100.00
	(1518)	(1407)	(1001)	(111)

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Business Register

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 15: Industry Comparison: Multi-unit

	NAICS agree 2 digit pct/obs	NAICS agree 3 digit pct/obs	NAICS agree 4 digit pct/obs	NAICS agree 5 digit pct/obs
No	27.30	33.40	43.98	52.69
Yes	72.70	66.60	56.02	47.31
Total	100.00 (1436)	100.00 (1410)	100.00 (914)	100.00 (186)

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Business Register

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 16: Single-unit Industry Comparison: BR NAICS when disagree with SIPP-EHC

	Single Units
	pct
Ag,Forest,FishHunt	0.53
Utilities	0.27
Construction	5.88
Manufacturing 1	2.67
Manufactuing 2	1.60
Manufacturing 3	2.94
Wholesale Trade	12.57
Retail Trade 1	3.74
Retail Trade 2	2.67
Transportation	1.34
Postal Service,Couriers,Warehousing	0.80
Information	2.67
Finance	1.07
Real Estate	4.55
Prof,Scientific,Technical Services	5.35
Admin.Support,Waste Management	13.90
Educational Services	2.94
Health Care,Social Assistance	2.67
Arts,Entertain,Rec	4.01
Accomodations,Food Service	4.28
Other Services	6.68
Public Admin	16.84
Total	100.00
<i>N</i>	374

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Business Register

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 17: Multi-unit Industry Comparison: BR NAICS when disagree with SIPP-EHC

	Multi Units
	pct
Mining,Oil,Gas	0.77
Utilities	0.26
Construction	1.53
Manufacturing 1	1.02
Manufactuing 2	3.06
Manufacturing 3	3.57
Wholesale Trade	9.44
Retail Trade 1	6.12
Retail Trade 2	4.08
Transportation	1.53
Postal Service,Couriers,Warehousing	2.81
Information	3.57
Finance	1.53
Real Estate	2.30
Prof,Scientific,Technical Services	4.08
Management of Companies	14.29
Admin.Support,Waste Management	12.24
Educational Services	4.08
Health Care,Social Assistance	10.46
Arts,Entertain,Rec	1.53
Accomodations,Food Service	8.16
Other Services	3.57
Total	100.00
<i>N</i>	392

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Business Register

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 18: NAICS Industries of SIPP-EHC Jobs for Respondents with no W-2s

	No W-2s pct
Ag,Forest,FishHunt	0.87
Construction	15.13
Manufacturing 1	2.09
Manufacturing 2	1.22
Manufacturing 3	2.26
Wholesale Trade	2.78
Retail Trade 1	6.43
Retail Trade 2	2.26
Transportation	7.13
Postal Service,Couriers,Warehousing	0.17
Information	2.09
Finance	1.74
Real Estate	3.30
Prof,Scientific,Technical Services	5.57
Admin.Support,Waste Management	8.70
Educational Services	3.48
Health Care,Social Assistance	10.96
Arts,Entertain,Rec	2.96
Accomodations,Food Service	7.48
Other Services	12.17
Public Admin	1.22
Total	100.00
<i>N</i>	575

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Business Register

Internal Revenue Service:

Form W-2 Records Tax Year 2009

Table 19: NAICS Industries of SIPP-EHC Jobs that do not match specific W-2 Records

	No W-2 match for SIPP job
	pct
Ag,Forest,FishHunt	0.73
Mining,Oil,Gas	0.28
Utilities	0.09
Construction	6.06
Manufacturing 1	1.93
Manufacturing 2	1.93
Manufacturing 3	3.58
Wholesale Trade	2.29
Retail Trade 1	7.16
Retail Trade 2	1.83
Transportation	3.58
Postal Service,Couriers,Warehousing	1.83
Information	2.39
Finance	2.94
Real Estate	3.03
Prof,Scientific,Technical Services	3.30
Admin.Support,Waste Management	7.80
Educational Services	9.36
Health Care,Social Assistance	15.23
Arts,Entertain,Rec	3.03
Accomodations,Food Service	7.16
Other Services	7.06
Public Admin	7.43
Total	100.00
<i>N</i>	1090

Sources:

U.S. Census Bureau:

Survey of Income and Program Participation Event History Calendar Field Test 2010

Business Register

Internal Revenue Service:

Form W-2 Records Tax Year 2009