Regional Economic Information System Data Compilation and Disclosure Estimation Process

This report describes processing steps implemented to prepare the U.S. Department of Commerce, Bureau of Economic Analysis, Regional Economic Information System employment and income data for use in the EPS-HDT application. The steps include downloading the data from the REIS web site, calculating estimates of employment and income when data are not disclosed by the Bureau of Economic Analysis, importing the reported and estimated data into a standard format used by the EPS-HDT application, and archiving previous years' data.

The purpose of this report is to: (1) Describe the process for downloading and organizing REIS data, (2) Describe the estimation method and the resulting accuracy, and (3) Provide a record for annual updates.

Contact:
Ray Rasker
406-570-7044
ray@headwaterseconomics.org

6/10/2010

Table of Contents

Introduction

This report describes processing steps implemented to prepare the U.S. Department of Commerce, Bureau of Economic Analysis, Regional Economic Information System employment and income data for use by the EPS-HDT application. The steps include downloading the data from the REIS web site, calculating estimates of employment and income when data are not disclosed by the Bureau of Economic Analysis, importing the reported and estimated data into a standard format used by the EPS-HDT application, and archiving previous years' data.

The purpose of this report is to: (1) Describe the process for downloading and organizing REIS data, (2) Describe the estimation method and the resulting accuracy, and (3) Provide a record for annual updates.

Estimation of Non-Disclosed REIS Employment and Income Numbers

When REIS does not report county level data for employment or income within the private sector, other REIS reported data from the same county and state are used to generate an estimate. In the REIS income tables (CA05 SIC and CA05N NAICS), when the BEA's non-disclosure flag specifies "Less than $50,000", we set the estimate to $25,000. In the employment tables (CA25 SIC and CA25N NAICS), when the BEA's non-disclosure flag specifies "Less than 10 jobs", we set the estimate to 5. For 45% of the remaining unreported records, we are able to generate estimates with consistently high accuracy using the method described in the following paragraphs. Due to a lack of information required to generate accurate estimates, 55% of unreported records are left without estimates.

Table 1. Counts of REIS records where data were estimated or left without estimates.

	Employment		Income	
	CA25N NAICS	CA25 SIC*	CA05N NAICS	CA05 SIC*
Total records (FIPS x Year x Industry Code)^	500472	1,115,670	500366	1021760
Reported as "Less than $50,000"/"Less than 10 jobs"	9215	24336	4118	13654
No reported data	125422	57,217	125422	56387
Total records estimated	46711	35,555	46711	35336
Total records left without an estimate	78711	21,662	78711	21051

* Estimates generated using 3 periods (1669-1976, 1977-1986, and 1987-2000) due to changes in SIC definitions.

^ Counts based on 2008 REIS county level data for the private sector.

The principle characteristic of REIS that enables us to generate estimates is that the data are represented in time series. For example, the CA25N table shows employment per county in many industries (such as retail trade, utilities, and manufacturing) from 2001 through the year of their latest release. When the data value for a year is not disclosed for a particular line code and county, but at least two other years of data are available, we generate estimates for year(s) for which data were not disclosed. For example, if employment was not reported for retail trade in Alachua County, FL for the years 2001 through 2005, but employment was reported for retail trade in Alachua County, FL in 2006

and 2007, we would use the method described below to generate estimates of employment in Alachua County, FL for the years 2001 through 2005.

For each year in a time series in which data are reported, the county's value is subtracted from the state's value multiplied by the share of the state's private employment that occurs within the county. Using our previous example, the reported employment in retail trade in Alachua County, FL would be subtracted from the reported employment in retail trade in the state of Florida in the corresponding year multiplied by the ratio of private employment in Alachua County over private employment in FL. This calculation is made for each year in which data are available (2006 and 2007 from our Alachua County example), and the results are averaged. This value is referred to as the "average difference". For each year in the time series in which data are not reported (2001 through 2005 in our Alachua County example), the estimate is generated by multiplying the "average difference" by the state's value times the share of the state's private employment that occurs within the county.

Seldom, the "average difference" is a negative number, which cannot occur for employment and is extremely rare for income figures. In these cases, the accuracy of our estimates was improved by using the following alternative approach. For each year in a time series in which data are reported, the county's value is divided by the state's value. Using our previous example, the reported employment in retail trade in Alachua County, FL would be divided by the reported employment in retail trade in the state of Florida. This calculation is made for each year in which data are available, and the results are averaged. This value is referred to as the "average proportion". For each year in which data are not reported and for which the "average difference" is negative, the estimate is generated by multiplying the "average proportion" by the state's value.

We conducted an accuracy assessment for all county/industry/year combinations where data were reported (365,835 records in CA25N; 1,034,117 records in CA25; 370,826 records in CA05N; and 951,719 records in CA05). The mean percent differences between the reported value and the estimated values are reported below.

	Employment		Income	
	CA25N NAICS	CA25 SIC*	CA05N NAICS	CA05 SIC*
Mean % Difference using "Average Difference"	-1.23%	-3.18%	-1.32%	-2.26%
Mean % Difference using "Average Proportion"	-1.62%	-4.08%	-4.18%	-8.51%

Details of the estimation methods are provided in Appendices B and C. Tables describing the accuracy assessment of estimates by industry are shown in Appendices D, E, F, and G.

Import the data from the web site

REIS data files are available on the internet at the following URL:

http://www.bea.gov/regional/docs/reis2007dvd.cfm

The files are comma delimited text files. REIS ca05, ca05n, ca25n and ca25 csv files were imported into ACCESS databases. The same process can be used with a MySQL database. The steps described below are implemented for each file.

The ca25n and ca05n files contain data for 2001-2008 using the NAICS classification. The following fields are included:

FIPS

Table – identifies which table ca25n (naics) or ca25 (sic)

Line code – using naics or sic classification

First year – the first year column heading

Line Title – description of the line code

Area Name – description for the FIPS code

Values by year – a column for each year with the reported employment or income. When data is not reported, there will be a "(L)" when data is set to 5 for employment or 25 for income, "(N)" when the data is not disclosed or "(D)" when the data is not available.

DISCL – includes a designation (0, 1, 2 or 9) for each year. This is a single field with the position corresponding to the year column. This field must be decomposed when normalizing the data.

The ca25 and ca05 files contain data for the years 1969-2000. The classification used during these years is the SIC classification but with enough differences during certain time periods that the data was sub-divided into three time periods for use in calculating estimates. The first time period is 1969-1976, the second is 1977-1986 and the third is 1987-2000. The data structure is the same as for the ca25n and ca05n except that there are 32 years worth of data. Each year is a field and the disclosure field has a value for each year (1 space for each year).

The following descriptions of the process use a single file as an example. File names can be changed.

Step 1: Import the csv file into ACCESS

Generally the wizard default settings are acceptable. The first line does contain field names. The year column data types must be set to "text' so the (N), (L) and (D) values are imported. This allows a cross check of the disclosure values when the data is normalized. Everything usually matches so technically the text disclosure values are not necessary as they will be set to null once the data has been normalized. An SQL script for appending the data is listed in Appendix A.

Step 2: Create the table REIS_emp

The REIS_emp table contains all the fields required to import the data from both the ca REIS data files as normalized data. There is a record for each year along with the associated disclosure value. There are additional fields that are used to set up the data required to calculate two methods for estimating employment data for records where the data was not disclosed (discl=1 or 2 or 9). The fields in the table are:

ID – an autonumber assigned to each record. This was used when randomly selecting 25% of the records for use in testing the chosen estimation method developed form the other 75% of the data.

Tbl – the REIS table name (ca05, ca05n, ca25 or ca25n – or any other ca table)

FIPStxt – the imported FIPS value

FipST – this is the state component of the "FIPStxt" field. This is required to assign the state values to each record.

FIPS – the numeric FIPS value used in EPS-HDT.

Line code – a text field with the line code assigned in REIS. This is related to the naics codes in a lookup table.

Linecode – numeric field for the final data tables; sometimes required for linking to other lookup tables which store the line codes as a numeric value.

NAICS/SIC – this is updated by linking to the lookup table "REIS_ca25linecodes" where the private employment/income line codes are assigned to either a NAICS or SIC classification for private employment or income. These records will be used to calculate estimated values when the data is not reported.

YR – the year for each record.

Emp – the reported employment for each fips by line code by year. When data is not disclosed there is a (D), (L) or (N) in this field. A 0 represents an actual reported value.

Discl – this has a single value (0, 1, 2 or 9) that identifies if the data is reported (0) or not. The correct column must be extracted from the ca** data when appending to the normalized table.

DisclVal – the numeric disclosure (Discl) value

REIS0090 – this is the total reported private employment for the fips by year.

STREIS0090 – this is the total reported private employment for the state fips by year.

STVal – this is the reported or estimated value for the corresponding state line code data record.

Best_Emp – the reported value when available or the best estimate using the STDiffCalc value when available and a positive value or the LCProp value when available. The remainder value is calculated from the values in this field. Initially it is set to the reported value but can be updated and the remainder query can be re-run if desired.

STDiffcalc – this is the calculated estimate using the average state difference method which is:

([ca].[stval]*([ca].[reis0090]/[ca].[streis0090]))+[avgSTdiff]

Where the average state difference is:

AvgSTDiff: Avg(([val]-([StVal]*([reis0090]/[streis0090]))))

LCProp – the calculated estimate using the state value times the average proportion for the FIPS by line code to the state value:

[ca].[stval]*[AvgPlineCode]

Where the average PlineCode - AvgPlinecode:

Avg([val]/[stval])

Step 3: Append the annual ca data

Append the data one year at a time into the "REIS_emp" table. For each year the year and appropriate column of the "discl" field must be appended. See Appendix A for the SQL scripts used to append the data. The ca05 file was too large to append into a single ACCESS database so it was sub-divided into the three time periods described above. Estimations were calculated within each time period.

Step 4: update the NAICS or SIC field

Use the lookup table REIS_ca25 linecodes to set the naics/sic value that matches the line code. Estimates will be done using the data set that has a value in the naics/sic field.

UPDATE REIS_emp INNER JOIN Reis_ca25linecodes

```
ON (REIS_emp.Code_type = Reis_ca25linecodes.type)
AND (REIS_emp.[Line code] = Reis_ca25linecodes.linecode)
SET REIS_emp.naics = [naics/sic];
```

Step 5: update the "emp" field

Run an update query to set the emp field to null where the field value is 1 or 9. Where discl=2 the VAL is set to 5 for the employment data (ca25 and ca25n) and 25 for the income data (ca05 and ca05n).

```
UPDATE REIS_emp SET REIS_emp.Emp = Null
WHERE (((REIS_emp.DISCL)="1"
Or (REIS_emp.DISCL)="9"));

UPDATE REIS_emp SET REIS_emp.Emp = 5    ------ for the income tables the estimate is set to 25
WHERE (REIS_emp.DISCL)="2";
```

Step 6: update REIS0090

Set the total reported value for the FIPS by Year (REIS0090) for each record – this helps to make the calculations faster although the same task can be accomplished using a sub-query.

```
UPDATE REIS_emp AS REIS_emp_1
INNER JOIN REIS_emp
ON (REIS_emp_1.FIPS = REIS_emp.FIPS)
AND (REIS_emp_1.YR = REIS_emp.YR)
SET REIS_emp_1.REIS0090 = [REIS_emp].[emp]
WHERE (((REIS_emp_1.naics) Is Not Null)
AND ((REIS_emp.[Line code])="0090"));
```

Step 7: update STVAL

Select the state total reported value from the table where the fips code is a state code and update the STVal field for every record. This makes estimate calculations faster but could be run as a sub-query.

```
UPDATE REIS_emp INNER JOIN REIS_emp AS REIS_emp_1
ON (REIS_emp.FipST = REIS_emp_1.FipST)
AND (REIS_emp.naics = REIS_emp_1.naics)
AND (REIS_emp.YR = REIS_emp_1.YR)
SET REIS_emp_1.STVal = [reis_emp].[Val]
WHERE (((Right([reis_emp].[fips],3))="000"));
```

Step 8: Make the Best_est table with the calculated estimates

Run the query "QryMaketblBestEst" to get all data used in making the estimates. This table is used to update the estimates for the final tables used in the EPS-HDT application. This query selects fields from the query "QryBestEst".

Step 12: Create the final tables for EPS-HDT

The tables for EPS-HDT have the following fields:

#FIPS – the numeric fips code – part of the primary key
#Linecode – the numeric line code value – part of the primary key
#YR – the numeric year – part of the primary key
VAL – this is the best estimate when there is no reported value
Discl – 0 for reported value, 1 for estimated value, 2 for forced (eg set to 5 or 25), 3 for not available
Tbl – the table name for the original data

The primary key is FIPS by Linecode by YR.

Step 13: Append metro/non-metro data

Data for the metro and non-metro areas are compiled to speed the retrieval of this information. Each FIPS code is assigned to the metro or the non-metro group. The following queries are used to append this summary data into the final data tables used by EPS-HDT for each REIS data set.

Several queries are run to compile the metro/non-metro data:

"QryMetroNM data": append the county level data to the ca table.

```
INSERT INTO ca05n_08 ( tbl, FIPS, Linecode, YR, Val, DISCL, naics )
SELECT ca05_emp.tbl, IIf([msa]="metro",Val("-" & Format(Left([fipstxt],2),"00") & "998"),IIf([msa]="non-
metro",Val("-" & Format(Left([fipstxt],2),"00") & "997"))) AS FIPS, ca05_emp.[Line code], ca05_emp.YR,
Sum(ca05_emp.Best_emp) AS SumOfBest_emp, Max(ca05_emp.DISCL) AS MaxOfDISCL, ca05_emp.naics
FROM ca05_emp INNER JOIN Georef02 ON ca05_emp.FIPS = Georef02.fips
GROUP BY ca05_emp.tbl, IIf([msa]="metro",Val("-" & Format(Left([fipstxt],2),"00") &
"998"),IIf([msa]="non-metro",Val("-" & Format(Left([fipstxt],2),"00") & "997"))), ca05_emp.[Line code],
ca05_emp.YR, ca05_emp.naics, Georef02.msa, Left([FIPStxt],2)
HAVING (((Georef02.msa)="metro" Or (Georef02.msa)="non-metro"));
```

"QryMetroUS": append the national level data to the ca table.

```
INSERT INTO ca05n_08 ( tbl, YR, fips, Linecode, Val, DISCL, naics )
SELECT ca05_emp.tbl, ca05_emp.YR, (IIf([msa]="metro",Val("-" & "998"),IIf([msa]="non-metro",Val("-" &
"997")))) AS fips, ca05_emp.[Line code], Sum(ca05_emp.Best_emp) AS SumOfBest_emp,
Max(ca05_emp.Discl) AS MaxOfDiscl, ca05_emp.naics
FROM ca05_emp INNER JOIN Georef02 ON ca05_emp.FIPS = Georef02.fips
```

```
GROUP BY ca05_emp.tbl, ca05_emp.YR, (IIf([msa]="metro",Val("-" & "998"),IIf([msa]="non-
metro",Val("-" & "997")))), ca05_emp.[Line code], ca05_emp.naics, Georef02.msa
HAVING (((Georef02.msa)="metro" Or (Georef02.msa)="non-metro"));
```

"QryMetroCalcs_crosstab": create a crosstab of the metro/non-metro data to be used for updating specific linecode calculations.

```
TRANSFORM Sum(ca05n_08.Val) AS SumOfVal
SELECT ca05n_08.tbl, ca05n_08.FIPS, ca05n_08.YR
FROM ca05n_08
WHERE (((ca05n_08.FIPS)<0))
GROUP BY ca05n_08.tbl, ca05n_08.FIPS, ca05n_08.YR
PIVOT ca05n_08.linecode;
```

"QryMaketblMnMCrossTab":

Create the table "ca05nMnMCrossTab" for the ca05 data to list each linecode value as a column. This table is used to calculate correct metro value for linecode=30 for the ca05 data. The value for linecode 30 is updated to the value of linecode 10 divided by the value of linecode 20.

```
SELECT QryMetroCalcs_crosstab.*
INTO ca05nMnMCrossTab
FROM QryMetroCalcs_crosstab;
```

Appendix A: Append annual ca05 data scripts

```
Dim str1969 As String
Dim str1970 As String
Dim str1971 As String
Dim str1972 As String
Dim str1973 As String
Dim str1974 As String
Dim str1975 As String
Dim str1976 As String
Dim str1977 As String
Dim str1978 As String
Dim str1979 As String
Dim str1980 As String
Dim str1981 As String
Dim str1982 As String
Dim str1983 As String
Dim str1984 As String
Dim str1985 As String
Dim str1986 As String
Dim str1987 As String
Dim str1988 As String
Dim str1989 As String
Dim str1990 As String
Dim str1991 As String
Dim str1992 As String
Dim str1993 As String
Dim str1994 As String
Dim str1995 As String
Dim str1996 As String
Dim str1997 As String
Dim str1998 As String
Dim str1999 As String
Dim str2000 As String

str1969 = "INSERT INTO ca05 ( tbl,FIPStxt, FIPS, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], [First year], [1969], Left([DISCL],1) " & _
" FROM [Reis_ca05txt];"

str1970 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1970', [1970], Mid([DISCL],2,1) " & _
" FROM [Reis_ca05txt];"

str1971 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1971', [1971], Mid([DISCL],3,1) " & _
```

```
" FROM [Reis_ca05txt];"

str1972 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1972', [1972], Mid([DISCL],4,1) " & _
" FROM [Reis_ca05txt];"

str1973 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1973', [1973], Mid([DISCL],5,1) " & _
" FROM [Reis_ca05txt];"

str1974 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1974', [1974], Mid([DISCL],6,1) " & _
" FROM [Reis_ca05txt];"

str1975 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1975', [1975], Mid([DISCL],7,1) " & _
" FROM [Reis_ca05txt];"

str1976 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1976', [1976], Mid([DISCL],8,1) " & _
" FROM [Reis_ca05txt];"

str1977 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1977', [1977], Mid([DISCL],9,1) " & _
" FROM [Reis_ca05txt];"

str1978 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1978', [1978], Mid([DISCL],10,1) " & _
" FROM [Reis_ca05txt];"

str1979 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1979', [1979], Mid([DISCL],11,1) " & _
" FROM [Reis_ca05txt];"

str1980 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1980', [1980], Mid([DISCL],12,1) " & _
" FROM [Reis_ca05txt];"

str1981 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1981', [1981], Mid([DISCL],13,1) " & _
"FROM [Reis_ca05txt];"

str1982 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
"SELECT table, FIPS, FIPS, [Line code], '1982', [1982], Mid([DISCL],14,1) " & _
" FROM [Reis_ca05txt];"

str1983 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
```

```
" SELECT table, FIPS, FIPS, [Line code], '1983', [1983], Mid([DISCL],15,1) " & _
" FROM [Reis_ca05txt];"

str1984 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1984', [1984], Mid([DISCL],16,1) " & _
" FROM [Reis_ca05txt];"

str1985 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1985', [1985], Mid([DISCL],17,1) " & _
" FROM [Reis_ca05txt];"

str1986 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1986', [1986], Mid([DISCL],18,1) " & _
" FROM [Reis_ca05txt];"

str1987 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1987', [1987], Mid([DISCL],19,1) " & _
" FROM [Reis_ca05txt];"

str1988 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1988', [1988], Mid([DISCL],20,1) " & _
" FROM [Reis_ca05txt];"

str1989 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1989', [1989], Mid([DISCL],21,1) " & _
" FROM [Reis_ca05txt];"

str1990 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1990', [1990], Mid([DISCL],22,1) " & _
" FROM [Reis_ca05txt];"

str1991 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1991', [1991], Mid([DISCL],23,1) " & _
" FROM [Reis_ca05txt];"

str1992 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1992', [1992], Mid([DISCL],24,1) " & _
" FROM [Reis_ca05txt];"

str1993 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1993', [1993], Mid([DISCL],25,1) " & _
" FROM [Reis_ca05txt];"

str1994 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1994', [1994], Mid([DISCL],26,1) " & _
" FROM [Reis_ca05txt];"
```

```
str1995 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1995', [1995], Mid([DISCL],27,1) " & _
" FROM [Reis_ca05txt];"

str1996 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1996', [1996], Mid([DISCL],28,1) " & _
" FROM [Reis_ca05txt];"

str1997 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1997', [1997], Mid([DISCL],29,1) " & _
" FROM [Reis_ca05txt];"

str1998 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1998', [1998], Mid([DISCL],30,1) " & _
" FROM [Reis_ca05txt];"

str1999 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '1999', [1999], Mid([DISCL],31,1) " & _
" FROM [Reis_ca05txt];"

str2000 = "INSERT INTO ca05( tbl, FIPS, FIPStxt, [Line code], YR, Emp, DISCL )" & _
" SELECT table, FIPS, FIPS, [Line code], '2000', [2000], Mid([DISCL],32,1) " & _
" FROM [Reis_ca05txt];"
DoCmd.SetWarnings False

DoCmd.RunSQL str1969
DoCmd.RunSQL str1970
DoCmd.RunSQL str1971
DoCmd.RunSQL str1972
DoCmd.RunSQL str1973
DoCmd.RunSQL str1974
DoCmd.RunSQL str1975
DoCmd.RunSQL str1976
DoCmd.RunSQL str1977
DoCmd.RunSQL str1978
DoCmd.RunSQL str1979
DoCmd.RunSQL str1980
DoCmd.RunSQL str1981
DoCmd.RunSQL str1982
DoCmd.RunSQL str1983
DoCmd.RunSQL str1984
DoCmd.RunSQL str1985
DoCmd.RunSQL str1986
DoCmd.RunSQL str1987
DoCmd.RunSQL str1988
DoCmd.RunSQL str1989
DoCmd.RunSQL str1990
```

```
DoCmd.RunSQL str1991
DoCmd.RunSQL str1992
DoCmd.RunSQL str1993
DoCmd.RunSQL str1994
DoCmd.RunSQL str1995
DoCmd.RunSQL str1996
DoCmd.RunSQL str1997
DoCmd.RunSQL str1998
DoCmd.RunSQL str1999
DoCmd.RunSQL str2000

DoCmd.SetWarnings True
```

Appendix B: QryCalcAvgDiff

```
SELECT ca.sicgroup,
ca.FIPS,
ca.[Line code],
ca.naics,
Avg(ca.Val) AS AvgOfEmp,
Avg(ca.StVal) AS AvgOfStVal,
Avg(([val]-([StVal]*([reis0090]/[streis0090])))) AS AvgSTDiff,
Avg([val]/[stval]) AS AvgPlinecode,
Count(ca.fips) AS CountOfRecs
FROM ca
WHERE (((ca.[Line code])<>"0090")
AND ((ca.val) Is Not Null)
AND ((ca.NAICS) Is Not Null))
GROUP BY ca.sicgroup,
ca.FIPS,
ca.[Line code],
ca.naics
HAVING (((Count(ca.fips))>1));
```

Appendix C: QryMaketblBestEst

Create the BestEst table – QryMaketblBestVal. It includes a field for the SICgrp for ca05 and ca25.

This query is based on 'QryBestEst":

```
SELECT ca.sicgroup,
ca.tbl,
ca.FIPSt,
ca.YR,
ca.FIPS,
ca.[Line code],
ca.naics,

The best value is the reported VAL when there is a reported value. When there is no reported value:
IIf(IsNull([val]),

        If the STCalcDiff  or average state  difference value is null:
        IIf(IsNull(([ca].[stval]*([ca].[reis0090]/[ca].[streis0090]))+[avgSTdiff]),

        If the LCProp or Line code proportion value is null:
        IIf(IsNull([ca].[StVal]*[AvgPlineCode]),Null,[ca].[stval]*[AvgPlineCode]),

        If the average state difference value is less than 0:
        IIf(([ca].[stval]*([ca].[reis0090]/[ca].[streis0090]))+[avgSTdiff]<0,

        Use the LCProp estimate when there is no reported value and the average State difference estimate is either null or negative:
        [ca].[stval]*[AvgPlineCode],

        Use the STcalcDiff estimate when there is no reported value and there is a >=0 average state difference calculated estimate:
        ([ca].[stval]*([ca].[reis0090]/[ca].[streis0090]))+[avgSTdiff])),

        Use the reported Val when it is available:
        [val]) AS BestEmp,
ca.Val,
ca.DISCL,
([ca].[stval]*([ca].[reis0090]/[ca].[streis0090]))+[avgSTdiff] AS STCalcDiff,
QryCalcAvgDiff.AvgOfEmp,
[ca].[StVal]*[AvgPlineCode] AS LCProp,
[ca].[stval]/[ca].[STReis0090] AS STProp,
ca.StVal,
```

```
QryCalcAvgDiff.AvgOfStVal,
QryCalcRemainder.CalcRemainder AS FIPSRemainder,
QryCalcAvgDiff.AvgPlinecode,
ca.REIS0090,
ca.STReis0090,
[calcremainder]/[ca].[reis0090] AS Remainderpercent,
QryFIPSxYRremainderSum.SumBestCalcDiff,
(([ca].[stval]/[ca].[STReis0090])/[sumBestCalcDiff])*[calcremainder] AS PofRemainder,
QryFIPSxYRremainderSum.[Count NullEmp]
FROM ((ca LEFT JOIN QryCalcAvgDiff
ON (ca.[Line code] = QryCalcAvgDiff.[Line code])
AND (ca.FIPS = QryCalcAvgDiff.FIPS)
AND (ca.sicgroup = QryCalcAvgDiff.sicgroup))
LEFT JOIN QryCalcRemainder
ON (ca.YR = QryCalcRemainder.YR)
AND (ca.FIPS = QryCalcRemainder.FIPS)
AND (ca.sicgroup = QryCalcRemainder.sicgroup))
LEFT JOIN QryFIPSxYRremainderSum
ON (ca.YR = QryFIPSxYRremainderSum.YR)
AND (ca.FIPS = QryFIPSxYRremainderSum.FIPS)
AND (ca.sicgroup = QryFIPSxYRremainderSum.sicgroup)
WHERE (((ca.sicgroup)=1)
AND ((ca.naics) Is Not Null))
ORDER BY ca.YR,
ca.FIPS,
ca.naics,
ca.Val;
```

Appendix D: Estimation Accuracy for CA25N (NAICS)

Table 1. Accuracy of REIS employment estimates (CA25N NAICS) using "Average Difference" method.

Line code	Description	Mean	Mean Estimate	Mean Difference	Mean Percent Difference	Count
100	Forestry, fishing, related activities, and other	424	429	0.0000	-1.18%	9875
200	Mining	465	471	0.0200	-2.35%	12930
300	Utilities	303	305	0.0175	-2.10%	11064
400	Construction	3708	3715	0.0000	-0.89%	22349
500	Manufacturing	5296	5322	0.0017	-1.80%	22566
600	Wholesale trade	2470	2480	0.0004	-1.73%	19527
700	Retail Trade	6072	6074	-0.0002	-0.31%	24424
800	Transportation and warehousing	2409	2434	0.0000	-1.47%	15912
900	Information	1404	1414	0.0012	-1.76%	20452
1000	Finance and insurance	3036	3045	0.0013	-0.44%	21640
1100	Real estate and rental and leasing	2494	2502	0.0002	-0.26%	21666
1200	Professional and technical services	4772	4812	0.0000	-0.60%	17894
1300	Management of companies and enterprises	1568	1590	0.1434	-8.13%	9035
1400	Administrative and waste services	4634	4678	0.0007	-1.78%	17400
1500	Educational services	1610	1624	0.0032	-1.34%	16362
1600	Health care and social assistance	8071	8155	0.0002	-0.31%	15665
1700	Arts, entertainment, and recreation	1397	1403	0.0004	-1.11%	19484
1800	Accommodation and food services	4663	4691	0.0000	-0.50%	19501
1900	Other services, except public administration	3379	3392	0.0000	-0.18%	22906

Table 1 continued. Accuracy of REIS employment estimates (CA25N NAICS) using "Average Difference" method.

		Percent of records in which the difference between the reported value and the estimate is within:					
Line code	Description	±1	±10	±50	±100	±500	> ±500
100	Forestry, fishing, related activities, and other	6.4%	46.4%	84.2%	90.8%	95.9%	1.0%
200	Mining	13.2%	50.0%	76.6%	85.5%	96.1%	2.3%
300	Utilities	17.3%	63.4%	88.1%	93.1%	98.0%	0.4%
400	Construction	2.2%	19.0%	54.9%	72.2%	94.3%	5.4%
500	Manufacturing	1.9%	14.1%	41.5%	57.6%	89.0%	10.5%
600	Wholesale trade	4.2%	29.6%	68.1%	81.0%	96.1%	3.4%
700	Retail Trade	2.2%	19.5%	55.8%	72.3%	94.7%	5.3%
800	Transportation and warehousing	3.3%	27.0%	63.0%	76.1%	94.0%	4.7%
900	Information	6.4%	43.0%	76.7%	85.6%	95.9%	3.1%
1000	Finance and insurance	4.1%	31.5%	72.5%	83.1%	95.7%	4.0%
1100	Real estate and rental and leasing	4.5%	30.6%	69.4%	82.1%	96.2%	3.5%
1200	Professional and technical services	4.4%	29.8%	64.0%	75.6%	92.9%	6.2%
1300	Management of companies and enterprises	3.4%	23.2%	56.4%	70.5%	91.5%	7.1%
1400	Administrative and waste services	3.4%	20.9%	50.5%	64.7%	90.1%	8.9%
1500	Educational services	13.0%	47.4%	75.5%	85.6%	96.8%	2.3%
1600	Health care and social assistance	2.4%	18.3%	49.2%	63.9%	90.4%	8.5%
1700	Arts, entertainment, and recreation	8.2%	41.4%	78.2%	88.2%	98.0%	1.5%
1800	Accommodation and food services	2.2%	19.5%	55.9%	73.2%	95.1%	4.2%
1900	Other services, except public administration	3.8%	31.4%	72.8%	85.3%	97.4%	2.1%

Table 2. Accuracy of REIS employment estimates (CA25N NAICS) using "Average Proportion" method.

Line code	Description	Mean	Mean Estimate	Mean Difference	Mean Percent Difference	Count
100	Forestry, fishing, related activities, and other	424	429	-0.0449	-1.32%	9875
200	Mining	465	471	0.3160	-4.32%	12930
300	Utilities	303	305	0.0419	-2.21%	11064
400	Construction	3708	3715	0.0182	-0.96%	22349
500	Manufacturing	5296	5322	0.0468	-2.13%	22566
600	Wholesale trade	2470	2480	-0.0355	-1.97%	19527
700	Retail Trade	6072	6074	-0.0064	-0.40%	24424
800	Transportation and warehousing	2409	2434	-0.0833	-1.62%	15912
900	Information	1404	1414	0.0557	-2.25%	20452
1000	Finance and insurance	3036	3045	0.0070	-0.71%	21640
1100	Real estate and rental and leasing	2494	2502	-0.0424	-0.71%	21666
1200	Professional and technical services	4772	4813	-0.1528	-1.01%	17894
1300	Management of companies and enterprises	1568	1590	0.5757	-9.40%	9035
1400	Administrative and waste services	4634	4679	-0.2965	-2.19%	17400
1500	Educational services	1610	1624	0.2015	-2.64%	16362
1600	Health care and social assistance	8071	8155	0.5340	-0.58%	15665
1700	Arts, entertainment, and recreation	1397	1403	-0.0112	-1.39%	19484
1800	Accommodation and food services	4663	4691	0.0856	-0.72%	19501
1900	Other services, except public administration	3379	3392	-0.0114	-0.22%	22906

Table 2 continued. Accuracy of REIS employment estimates (CA25N NAICS) using "Average Proportion" method.

		Percent of records in which the difference between the reported value and the estimate is within:					
Line code	Description	±1	±10	±50	±100	±500	> ±500
100	Forestry, fishing, related activities, and other	6.9%	49.9%	87.0%	92.8%	96.3%	0.7%
200	Mining	14.8%	55.1%	81.8%	90.0%	97.5%	0.9%
300	Utilities	20.6%	64.6%	88.0%	93.3%	97.9%	0.5%
400	Construction	2.5%	19.4%	55.6%	72.2%	93.9%	5.8%
500	Manufacturing	2.4%	14.5%	40.3%	55.6%	87.8%	11.6%
600	Wholesale trade	4.8%	31.0%	68.4%	80.4%	95.4%	4.1%
700	Retail Trade	2.4%	20.7%	56.5%	71.7%	93.1%	6.8%
800	Transportation and warehousing	3.9%	27.6%	63.9%	76.5%	93.9%	4.8%
900	Information	9.3%	49.0%	78.5%	86.3%	95.8%	3.2%
1000	Finance and insurance	5.0%	35.9%	73.8%	82.9%	94.9%	4.8%
1100	Real estate and rental and leasing	5.7%	36.1%	74.2%	84.8%	96.8%	2.9%
1200	Professional and technical services	5.6%	35.9%	67.5%	77.2%	92.6%	6.5%
1300	Management of companies and enterprises	5.4%	30.1%	60.9%	73.3%	92.0%	6.6%
1400	Administrative and waste services	3.6%	21.9%	50.2%	63.6%	88.9%	10.1%
1500	Educational services	25.0%	52.8%	77.0%	85.8%	96.8%	2.3%
1600	Health care and social assistance	3.5%	20.2%	50.7%	64.8%	90.1%	8.8%
1700	Arts, entertainment, and recreation	10.4%	44.6%	78.8%	88.1%	98.1%	1.4%
1800	Accommodation and food services	2.3%	19.4%	55.0%	71.8%	93.5%	5.8%
1900	Other services, except public administration	4.5%	32.6%	72.9%	84.5%	97.1%	2.4%

Appendix E: Estimation Accuracy for CA05N (NAICS)

Table 1. Accuracy of REIS income estimates (CA05N NAICS) using "Average Difference" method.

Line code	Description	Mean	Mean Estimate	Mean Difference	Mean Percent Difference	Count
100	Forestry, fishing, related activities, and other	$11,979	$12,123	$0.00	-1.94%	9867
200	Mining	$44,275	$44,863	$0.08	21.41%	12803
300	Utilities	$30,580	$30,854	-$0.13	-3.67%	12181
400	Construction	$186,620	$186,935	$0.02	-2.20%	22344
500	Manufacturing	$332,357	$333,974	-$0.07	-1.04%	22526
600	Wholesale trade	$163,162	$163,823	-$0.03	-2.36%	19498
700	Retail Trade	$173,439	$173,495	-$0.02	-0.71%	24424
800	Transportation and warehousing	$119,481	$120,776	$0.00	-2.39%	15911
900	Information	$109,402	$110,292	$0.11	-4.46%	20370
1000	Finance and insurance	$217,405	$218,074	$0.02	0.04%	21618
1100	Real estate and rental and leasing	$63,685	$63,880	$0.01	-3.61%	21639
1200	Professional and technical services	$315,989	$318,740	$0.00	0.73%	17893
1300	Management of companies and enterprises	$164,224	$166,453	-$2.56	-11.65%	8514
1400	Administrative and waste services	$138,502	$139,862	$0.05	-3.47%	17385
1500	Educational services	$52,129	$52,581	$0.11	-1.77%	16076
1600	Health care and social assistance	$371,665	$375,686	$0.01	-0.67%	15650
1700	Arts, entertainment, and recreation	$36,114	$36,286	$0.03	-5.04%	19448
1800	Accommodation and food services	$96,150	$96,784	$0.00	-0.69%	19493
1900	Other services, except public administration	$102,708	$103,117	$0.00	-0.78%	22906

Table 1 continued. Accuracy of REIS income estimates (CA05N NAICS) using "Average Difference" method.

		Percent of records in which the difference between the reported value and the estimate is within:					
Line code	Description	±$50	±$100	±$500	±$1000	±$5000	±$10000
100	Forestry, fishing, related activities, and other	11.0%	20.4%	56.6%	73.4%	91.5%	94.0%
200	Mining	15.0%	22.3%	43.2%	54.0%	77.6%	85.4%
300	Utilities	15.6%	23.7%	50.4%	63.5%	86.0%	91.6%
400	Construction	2.7%	5.2%	20.5%	33.3%	69.8%	81.6%
500	Manufacturing	2.1%	3.7%	14.0%	23.5%	55.9%	71.4%
600	Wholesale trade	4.1%	7.6%	28.2%	42.8%	76.1%	85.3%
700	Retail Trade	4.4%	8.2%	29.6%	44.4%	78.4%	87.8%
800	Transportation and warehousing	3.9%	7.3%	26.8%	40.7%	73.6%	83.6%
900	Information	7.4%	13.6%	42.6%	57.5%	82.8%	88.8%
1000	Finance and insurance	3.8%	7.4%	27.9%	42.6%	74.4%	83.4%
1100	Real estate and rental and leasing	8.5%	15.0%	44.4%	60.7%	86.8%	92.3%
1200	Professional and technical services	3.5%	6.8%	23.9%	36.7%	67.0%	77.2%
1300	Management of companies and enterprises	1.2%	2.7%	11.9%	20.2%	53.4%	67.5%
1400	Administrative and waste services	5.4%	9.3%	27.8%	41.4%	73.6%	84.4%
1500	Educational services	16.9%	26.1%	52.1%	63.6%	86.9%	92.4%
1600	Health care and social assistance	3.2%	6.0%	21.5%	32.8%	62.3%	74.2%
1700	Arts, entertainment, and recreation	16.2%	25.6%	57.2%	70.8%	90.3%	94.4%
1800	Accommodation and food services	7.0%	12.6%	40.2%	56.0%	84.7%	91.5%
1900	Other services, except public administration	5.2%	9.4%	33.9%	50.8%	84.7%	92.1%

Table 2. Accuracy of REIS income estimates (CA05N NAICS) using "Average Proportion" method.

Line code	Description	Mean	Mean Estimate	Mean Difference	Mean Percent Difference	Count
100	Forestry, fishing, related activities, and other	$11,979	$12,127	-$4.14	-2.82%	9867
200	Mining	$44,275	$44,824	$39.17	-13.68%	12803
300	Utilities	$30,580	$30,822	$32.67	-9.50%	12181
400	Construction	$186,620	$186,934	$1.69	-2.75%	22344
500	Manufacturing	$332,357	$333,972	$1.98	-3.63%	22526
600	Wholesale trade	$163,162	$163,834	-$11.24	-3.73%	19498
700	Retail Trade	$173,439	$173,496	-$1.38	-0.71%	24424
800	Transportation and warehousing	$119,481	$120,786	-$10.07	-2.69%	15911
900	Information	$109,402	$110,289	$3.14	-5.62%	20370
1000	Finance and insurance	$217,405	$218,074	$0.10	-1.64%	21618
1100	Real estate and rental and leasing	$63,685	$63,879	$0.58	-4.41%	21639
1200	Professional and technical services	$315,989	$318,753	-$13.43	-2.60%	17893
1300	Management of companies and enterprises	$164,224	$166,376	$74.47	-16.98%	8514
1400	Administrative and waste services	$138,502	$139,865	-$2.87	-5.73%	17385
1500	Educational services	$52,129	$52,571	$10.43	-8.02%	16076
1600	Health care and social assistance	$371,665	$375,664	$22.69	-1.14%	15650
1700	Arts, entertainment, and recreation	$36,114	$36,286	$0.05	-5.97%	19448
1800	Accommodation and food services	$96,150	$96,782	$2.46	-1.17%	19493
1900	Other services, except public administration	$102,708	$103,120	-$2.39	-0.52%	22906

Table 2 continued. Accuracy of REIS income estimates (CA05N NAICS) using "Average Proportion" method.

		Percent of records in which the difference between the reported value and the estimate is within:					
Line code	Description	±$50	±$100	±$500	±$1000	±$5000	±$10000
100	Forestry, fishing, related activities, and other	13.1%	23.5%	63.1%	79.0%	94.1%	95.8%
200	Mining	26.9%	34.7%	52.4%	63.2%	85.2%	91.3%
300	Utilities	23.6%	31.0%	57.0%	68.4%	88.5%	92.9%
400	Construction	3.1%	5.8%	22.3%	35.2%	70.8%	82.5%
500	Manufacturing	3.4%	5.6%	16.5%	25.4%	55.8%	70.1%
600	Wholesale trade	5.4%	10.0%	33.1%	48.8%	78.7%	86.6%
700	Retail Trade	5.2%	9.8%	33.8%	48.8%	80.5%	88.7%
800	Transportation and warehousing	4.6%	8.0%	29.0%	43.9%	76.3%	85.7%
900	Information	12.3%	21.5%	52.9%	65.3%	85.3%	90.1%
1000	Finance and insurance	6.5%	12.1%	40.7%	56.8%	82.3%	88.4%
1100	Real estate and rental and leasing	11.9%	19.5%	50.5%	64.5%	87.2%	92.3%
1200	Professional and technical services	8.6%	14.8%	40.6%	53.3%	76.5%	83.7%
1300	Management of companies and enterprises	3.9%	6.8%	22.0%	33.6%	63.3%	74.2%
1400	Administrative and waste services	8.6%	13.3%	33.2%	45.5%	74.4%	84.3%
1500	Educational services	33.7%	41.8%	64.3%	74.4%	90.6%	94.6%
1600	Health care and social assistance	5.3%	9.0%	28.4%	40.5%	68.3%	79.7%
1700	Arts, entertainment, and recreation	23.1%	34.1%	64.8%	75.9%	92.1%	95.6%
1800	Accommodation and food services	9.1%	15.4%	44.4%	60.0%	86.8%	92.6%
1900	Other services, except public administration	6.8%	12.8%	42.5%	59.7%	87.4%	93.2%

Appendix F: Estimation Accuracy for CA25 (SIC)

Table 1. Accuracy of REIS employment estimates (CA25 SIC) using "Average Difference" method.

Line code	Description	Years	Mean	Mean Estimate	Mean Difference	Mean Percent Difference	Count
100	Agricultural services, forestry, fishing & other	1969-1977	199	199	0.0019	-6.51%	23329
200	Mining	1969-1977	301	301	0.1921	-18.54%	20352
300	Construction	1969-1977	1535	1535	0.0000	-7.08%	24464
400	Manufacturing	1969-1977	6556	6556	0.0191	-3.62%	23983
500	Transportation and public utilities	1969-1977	1665	1665	0.0002	-2.02%	23655
610	Wholesale trade	1969-1977	1527	1527	0.0020	-6.73%	23440
620	Retail trade	1969-1977	4712	4712	0.0000	-0.41%	24792
700	Finance, insurance, and real estate	1969-1977	2234	2234	-0.0002	-1.28%	24374
800	Services	1969-1977	6150	6150	0.0000	-0.52%	24200
900	Government and government enterprises	1969-1977	5372	5372	0.0000	-0.36%	24856
100	Agricultural services, forestry, fishing & other	1977-1986	325	325	0.0027	-6.31%	29476
200	Mining	1977-1986	487	487	0.0797	-13.13%	25757
300	Construction	1977-1986	1884	1884	0.0000	-4.30%	30513
400	Manufacturing	1977-1986	6717	6718	0.0013	-2.57%	29939
500	Transportation and public utilities	1977-1986	1876	1877	0.0000	-1.58%	29800
610	Wholesale trade	1977-1986	1929	1929	0.0006	-2.24%	29580
620	Retail trade	1977-1986	5905	5905	0.0000	-0.35%	31007
700	Finance, insurance, and real estate	1977-1986	2937	2937	-0.0004	-0.67%	30395
800	Services	1977-1986	8779	8780	0.0000	-0.32%	30205
900	Government and government enterprises	1977-1986	6005	6005	0.0000	-0.34%	31047
100	Agricultural services, forestry, fishing & other	1987-2000	572	573	-0.0010	-4.15%	34431
200	Mining	1987-2000	400	401	0.0849	-9.85%	30708
300	Construction	1987-2000	2546	2546	0.0008	-3.03%	41675
400	Manufacturing	1987-2000	6427	6427	0.0072	-3.88%	41881
500	Transportation and public utilities	1987-2000	2308	2308	0.0000	-1.87%	41855
610	Wholesale trade	1987-2000	2320	2320	0.0059	-2.81%	41148
620	Retail trade	1987-2000	7733	7734	0.0000	-0.50%	43386
700	Finance, insurance, and real estate	1987-2000	3738	3739	0.0011	-1.15%	41785
800	Services	1987-2000	13966	13967	0.0000	-0.12%	42371
900	Government and government enterprises	1987-2000	6920	6920	0.0000	-0.65%	43523

Table 1 continued. Accuracy of REIS employment estimates (CA25 SIC) using "Average Difference" method.

Line code	Description	Years	Percent of records in which the difference between the reported value and the estimate is within: ±1	±10	±50	±100	±500	> ±500
100	Agricultural services, forestry, fishing & other	1969-1977	10.3%	56.3%	91.4%	96.3%	99.5%	0.5%
200	Mining	1969-1977	7.9%	45.8%	81.4%	90.2%	98.7%	1.3%
300	Construction	1969-1977	2.3%	20.1%	60.3%	76.9%	95.7%	4.3%
400	Manufacturing	1969-1977	1.4%	12.3%	39.9%	57.6%	90.3%	9.7%
500	Transportation and public utilities	1969-1977	3.8%	31.4%	75.2%	87.8%	98.3%	1.7%
610	Wholesale trade	1969-1977	3.0%	24.9%	68.7%	84.1%	97.6%	2.4%
620	Retail trade	1969-1977	1.8%	15.8%	54.5%	74.0%	96.1%	3.9%
700	Finance, insurance, and real estate	1969-1977	2.8%	24.2%	68.4%	83.7%	97.3%	2.7%
800	Services	1969-1977	1.6%	14.1%	48.8%	68.2%	94.1%	5.9%
900	Government and government enterprises	1969-1977	1.5%	13.6%	48.0%	66.8%	91.2%	8.8%
100	Agricultural services, forestry, fishing & other	1977-1986	7.3%	48.9%	86.2%	93.5%	99.2%	0.8%
200	Mining	1977-1986	8.3%	40.1%	71.9%	82.1%	96.4%	3.5%
300	Construction	1977-1986	2.2%	19.8%	57.7%	73.7%	94.2%	5.8%
400	Manufacturing	1977-1986	1.2%	10.7%	35.6%	52.5%	87.5%	12.5%
500	Transportation and public utilities	1977-1986	3.5%	27.5%	67.4%	81.3%	96.6%	3.4%
610	Wholesale trade	1977-1986	3.4%	26.3%	67.7%	82.3%	97.1%	2.9%
620	Retail trade	1977-1986	1.5%	14.7%	50.2%	69.7%	93.6%	6.4%
700	Finance, insurance, and real estate	1977-1986	2.9%	25.0%	65.6%	80.3%	95.7%	4.3%
800	Services	1977-1986	1.4%	12.8%	45.1%	63.8%	92.4%	7.6%
900	Government and government enterprises	1977-1986	1.3%	12.1%	43.1%	61.4%	90.1%	9.9%
100	Agricultural services, forestry, fishing & other	1987-2000	4.7%	36.9%	80.8%	90.0%	98.0%	1.7%
200	Mining	1987-2000	9.6%	44.1%	76.9%	87.0%	98.0%	1.7%
300	Construction	1987-2000	2.0%	17.6%	52.3%	69.8%	94.1%	5.9%
400	Manufacturing	1987-2000	1.1%	9.5%	31.4%	47.4%	84.6%	15.4%
500	Transportation and public utilities	1987-2000	2.6%	21.5%	59.3%	74.9%	94.5%	5.5%
610	Wholesale trade	1987-2000	2.7%	22.4%	61.5%	77.2%	95.3%	4.7%
620	Retail trade	1987-2000	1.3%	12.7%	44.4%	61.9%	90.3%	9.7%
700	Finance, insurance, and real estate	1987-2000	2.4%	20.5%	59.8%	75.1%	94.0%	6.0%
800	Services	1987-2000	0.9%	8.7%	33.4%	50.3%	85.6%	14.3%
900	Government and government enterprises	1987-2000	1.1%	10.1%	36.8%	54.2%	86.5%	13.5%

Table 2. Accuracy of REIS employment estimates (CA25 SIC) using "Average Proportion" method.

Line code	Description	Year	Mean	Mean Estimate	Mean Difference	Mean Percent Difference	Count
100	Agricultural services, forestry, fishing & other	1969-1977	199	199	0.0191	-6.69%	23329
200	Mining	1969-1977	301	301	0.2285	-19.79%	20352
300	Construction	1969-1977	1535	1535	-0.0875	-7.76%	24464
400	Manufacturing	1969-1977	6556	6556	0.0204	-5.57%	23983
500	Transportation and public utilities	1969-1977	1665	1665	-0.0033	-2.62%	23655
610	Wholesale trade	1969-1977	1527	1527	-0.0518	-8.96%	23440
620	Retail trade	1969-1977	4712	4712	0.0041	-0.99%	24792
700	Finance, insurance, and real estate	1969-1977	2234	2234	-0.0073	-2.58%	24374
800	Services	1969-1977	6150	6150	0.0461	-1.22%	24200
900	Government and government enterprises	1969-1977	5372	5372	-0.0002	-0.48%	24856
100	Agricultural services, forestry, fishing & other	1977-1986	325	325	0.0004	-6.14%	29476
200	Mining	1977-1986	487	487	0.2732	-17.00%	25757
300	Construction	1977-1986	1884	1884	0.0094	-5.07%	30513
400	Manufacturing	1977-1986	6717	6718	0.0063	-3.83%	29939
500	Transportation and public utilities	1977-1986	1876	1877	-0.0869	-2.28%	29800
610	Wholesale trade	1977-1986	1929	1929	-0.0079	-3.39%	29580
620	Retail trade	1977-1986	5905	5905	-0.0235	-0.85%	31007
700	Finance, insurance, and real estate	1977-1986	2937	2937	-0.0128	-1.21%	30395
800	Services	1977-1986	8779	8780	0.0400	-0.95%	30205
900	Government and government enterprises	1977-1986	6005	6005	-0.0199	-0.42%	31047
100	Agricultural services, forestry, fishing & other	1987-2000	572	573	0.0271	-3.73%	34431
200	Mining	1987-2000	400	401	0.2510	-12.28%	30708
300	Construction	1987-2000	2546	2546	0.0303	-3.45%	41675
400	Manufacturing	1987-2000	6427	6427	-0.0334	-4.88%	41881
500	Transportation and public utilities	1987-2000	2308	2308	0.1024	-3.27%	41855
610	Wholesale trade	1987-2000	2320	2320	0.0288	-4.24%	41148
620	Retail trade	1987-2000	7733	7734	-0.0215	-1.01%	43386
700	Finance, insurance, and real estate	1987-2000	3738	3739	-0.0357	-2.11%	41785
800	Services	1987-2000	13966	13967	0.1837	-1.30%	42371
900	Government and government enterprises	1987-2000	6920	6920	-0.0033	-0.73%	43523

Table 2 continued. Accuracy of REIS employment estimates (CA25 SIC) using "Average Proportion" method.

Line code	Description	Years	Percent of records in which the difference between the reported value and the estimate is within: ±1	±10	±50	±100	±500	> ±500
100	Agricultural services, forestry, fishing & other	1969-1977	11.7%	58.6%	92.5%	97.4%	99.9%	0.1%
200	Mining	1969-1977	10.2%	51.1%	85.6%	93.1%	99.3%	0.7%
300	Construction	1969-1977	2.3%	19.6%	58.4%	75.4%	95.3%	4.7%
400	Manufacturing	1969-1977	1.7%	13.2%	38.1%	54.2%	87.2%	12.8%
500	Transportation and public utilities	1969-1977	4.7%	34.0%	76.3%	87.6%	97.7%	2.3%
610	Wholesale trade	1969-1977	4.5%	28.9%	70.1%	84.1%	96.9%	3.1%
620	Retail trade	1969-1977	1.7%	14.8%	52.4%	72.9%	94.5%	5.5%
700	Finance, insurance, and real estate	1969-1977	3.2%	26.5%	71.5%	84.7%	96.5%	3.5%
800	Services	1969-1977	1.6%	13.8%	48.4%	67.5%	93.3%	6.7%
900	Government and government enterprises	1969-1977	2.2%	19.9%	59.5%	74.4%	91.8%	8.2%
100	Agricultural services, forestry, fishing & other	1977-1986	8.0%	50.6%	87.1%	93.9%	99.4%	0.6%
200	Mining	1977-1986	13.6%	48.2%	76.8%	85.6%	97.6%	2.3%
300	Construction	1977-1986	2.1%	18.8%	56.4%	72.0%	93.5%	6.4%
400	Manufacturing	1977-1986	1.8%	12.7%	35.2%	50.0%	84.2%	15.8%
500	Transportation and public utilities	1977-1986	4.2%	29.0%	68.3%	81.7%	96.1%	3.9%
610	Wholesale trade	1977-1986	4.7%	29.7%	69.6%	83.0%	96.3%	3.7%
620	Retail trade	1977-1986	1.5%	13.3%	47.4%	66.7%	91.7%	8.3%
700	Finance, insurance, and real estate	1977-1986	3.9%	29.2%	68.4%	81.4%	95.7%	4.3%
800	Services	1977-1986	1.4%	11.9%	42.2%	60.7%	89.5%	10.4%
900	Government and government enterprises	1977-1986	2.0%	18.3%	56.2%	72.3%	93.1%	6.9%
100	Agricultural services, forestry, fishing & other	1987-2000	5.0%	39.6%	82.1%	91.3%	98.6%	1.1%
200	Mining	1987-2000	15.8%	50.1%	81.2%	90.0%	98.6%	1.1%
300	Construction	1987-2000	2.1%	17.8%	52.4%	69.5%	93.3%	6.7%
400	Manufacturing	1987-2000	1.4%	10.9%	31.2%	45.5%	82.1%	17.9%
500	Transportation and public utilities	1987-2000	2.9%	22.5%	61.1%	76.1%	94.4%	5.6%
610	Wholesale trade	1987-2000	3.5%	24.8%	63.1%	77.3%	94.3%	5.7%
620	Retail trade	1987-2000	1.3%	11.6%	41.2%	58.4%	87.4%	12.6%
700	Finance, insurance, and real estate	1987-2000	2.7%	22.1%	60.1%	74.3%	92.4%	7.5%
800	Services	1987-2000	0.9%	8.0%	31.2%	48.9%	83.7%	16.3%
900	Government and government enterprises	1987-2000	1.4%	14.1%	45.9%	61.8%	88.2%	11.8%

Appendix F: Estimation Accuracy for CA05 (SIC)

Table 1. Accuracy of REIS income estimates (CA05 SIC) using "Average Difference" method.

Line code	Description	Years	Mean	Mean Estimate	Mean Difference	Mean Percent Difference	Count
100	Agricultural services, forestry, fishing & other	1969-1977	$1,387	$1,387	$0.00	-6.04%	23401
200	Mining	1969-1977	$4,036	$4,035	$0.74	30.70%	21004
300	Construction	1969-1977	$18,699	$18,699	$0.00	-9.36%	24464
400	Manufacturing	1969-1977	$74,550	$74,550	-$0.65	5.67%	23925
500	Transportation and public utilities	1969-1977	$21,740	$21,740	$0.00	0.26%	23649
610	Wholesale trade	1969-1977	$18,324	$18,324	$0.01	3.96%	23440
620	Retail trade	1969-1977	$30,242	$30,242	$0.00	-0.27%	24792
700	Finance, insurance, and real estate	1969-1977	$15,417	$15,417	$0.01	2.29%	24373
800	Services	1969-1977	$46,209	$46,209	$0.00	1.31%	24200
100	Agricultural services, forestry, fishing & other	1977-1986	$3,312	$3,313	$0.02	-11.59%	29487
200	Mining	1977-1986	$12,770	$12,778	$0.12	-10.76%	27547
300	Construction	1977-1986	$40,784	$40,786	$0.00	-7.54%	30513
400	Manufacturing	1977-1986	$159,497	$159,540	$0.17	2.32%	29901
500	Transportation and public utilities	1977-1986	$49,626	$49,634	$0.00	-0.90%	29800
610	Wholesale trade	1977-1986	$44,005	$44,015	$0.02	1.64%	29578
620	Retail trade	1977-1986	$64,447	$64,447	$0.00	-0.41%	31007
700	Finance, insurance, and real estate	1977-1986	$41,141	$41,144	-$0.01	1.91%	30391
800	Services	1977-1986	$132,285	$132,301	$0.00	1.86%	30205
100	Agricultural services, forestry, fishing & other	1987-2000	$9,106	$9,111	$0.02	-9.90%	34424
200	Mining	1987-2000	$19,194	$19,215	$3.88	-41.09%	28985
300	Construction	1987-2000	$87,839	$87,844	$0.01	-6.53%	41675
400	Manufacturing	1987-2000	$268,234	$268,246	$0.02	-6.94%	41703
500	Transportation and public utilities	1987-2000	$96,426	$96,435	$0.01	-1.69%	41850
610	Wholesale trade	1987-2000	$96,011	$96,026	$0.19	0.14%	41106
620	Retail trade	1987-2000	$134,154	$134,156	$0.00	-1.32%	43386
700	Finance, insurance, and real estate	1987-2000	$123,485	$123,512	$0.01	7.76%	41747
800	Services	1987-2000	$398,884	$398,911	$0.00	1.92%	42371

Table 1 continued. Accuracy of REIS income estimates (CA05 SIC) using "Average Difference" method.

Line code	Description	Years	Percent of records in which the difference between the reported value and the estimate is within: ±$50	±$100	±$500	±$1000	±$5000	±$10000
100	Agricultural services, forestry, fishing & other	1969-1977	44.6%	63.4%	91.4%	95.8%	99.3%	99.8%
200	Mining	1969-1977	24.2%	38.1%	70.3%	80.3%	94.8%	97.7%
300	Construction	1969-1977	10.6%	19.6%	55.2%	70.9%	92.9%	96.7%
400	Manufacturing	1969-1977	5.4%	10.2%	36.0%	52.4%	84.0%	91.4%
500	Transportation and public utilities	1969-1977	13.3%	24.0%	62.5%	77.0%	94.8%	97.5%
610	Wholesale trade	1969-1977	12.1%	21.7%	60.3%	76.3%	95.1%	97.7%
620	Retail trade	1969-1977	12.3%	22.2%	61.5%	77.4%	95.3%	98.0%
700	Finance, insurance, and real estate	1969-1977	24.7%	38.8%	73.5%	83.8%	96.0%	98.2%
800	Services	1969-1977	10.8%	19.4%	52.8%	68.3%	91.8%	95.7%
100	Agricultural services, forestry, fishing & other	1977-1986	24.3%	39.3%	78.0%	88.1%	97.8%	99.0%
200	Mining	1977-1986	16.7%	26.4%	56.2%	67.8%	88.4%	94.0%
300	Construction	1977-1986	5.8%	11.3%	39.1%	55.0%	84.3%	91.6%
400	Manufacturing	1977-1986	2.9%	5.7%	21.6%	33.9%	68.7%	81.3%
500	Transportation and public utilities	1977-1986	6.4%	12.3%	40.9%	57.4%	86.3%	92.8%
610	Wholesale trade	1977-1986	6.9%	13.2%	43.2%	60.4%	89.1%	94.3%
620	Retail trade	1977-1986	7.2%	13.4%	43.4%	59.7%	87.5%	93.6%
700	Finance, insurance, and real estate	1977-1986	9.8%	18.1%	49.5%	64.4%	88.6%	93.6%
800	Services	1977-1986	4.4%	8.3%	30.4%	45.1%	78.9%	88.4%
100	Agricultural services, forestry, fishing & other	1987-2000	13.8%	24.8%	63.0%	78.0%	95.1%	97.4%
200	Mining	1987-2000	10.5%	18.4%	46.2%	59.8%	85.7%	92.1%
300	Construction	1987-2000	3.2%	6.3%	24.3%	38.0%	73.8%	84.9%
400	Manufacturing	1987-2000	1.7%	3.4%	13.7%	23.0%	55.4%	70.2%
500	Transportation and public utilities	1987-2000	3.5%	6.7%	26.5%	41.9%	77.4%	86.6%
610	Wholesale trade	1987-2000	3.9%	7.5%	28.5%	43.7%	79.2%	88.4%
620	Retail trade	1987-2000	3.7%	7.1%	26.8%	41.8%	75.5%	85.6%
700	Finance, insurance, and real estate	1987-2000	4.5%	8.6%	29.8%	43.3%	74.4%	84.0%
800	Services	1987-2000	2.0%	4.0%	16.0%	26.7%	60.6%	74.4%

Table 2. Accuracy of REIS income estimates (CA05 SIC) using "Average Proportion" method.

Line code	Description	Year	Mean	Mean Estimate	Mean Difference	Mean Percent Difference	Count
100	Agricultural services, forestry, fishing & other	1969-1977	$1,387	$1,387	$0.31	-6.26%	23401
200	Mining	1969-1977	$4,036	$4,032	$3.28	-31.69%	21004
300	Construction	1969-1977	$18,699	$18,706	-$6.64	-10.98%	24464
400	Manufacturing	1969-1977	$74,550	$74,549	$0.22	-10.26%	23925
500	Transportation and public utilities	1969-1977	$21,740	$21,741	-$1.69	-3.11%	23649
610	Wholesale trade	1969-1977	$18,324	$18,325	-$0.70	-8.09%	23440
620	Retail trade	1969-1977	$30,242	$30,242	$0.09	-0.80%	24792
700	Finance, insurance, and real estate	1969-1977	$15,417	$15,418	-$0.08	-1.89%	24373
800	Services	1969-1977	$46,209	$46,209	$0.78	-1.32%	24200
100	Agricultural services, forestry, fishing & other	1977-1986	$3,312	$3,313	-$0.22	-10.61%	29487
200	Mining	1977-1986	$12,770	$12,770	$8.41	-25.81%	27547
300	Construction	1977-1986	$40,784	$40,785	$1.88	-9.12%	30513
400	Manufacturing	1977-1986	$159,497	$159,541	-$0.84	-7.15%	29901
500	Transportation and public utilities	1977-1986	$49,626	$49,656	-$21.41	-3.14%	29800
610	Wholesale trade	1977-1986	$44,005	$44,015	-$0.10	-5.30%	29578
620	Retail trade	1977-1986	$64,447	$64,448	-$0.76	-1.01%	31007
700	Finance, insurance, and real estate	1977-1986	$41,141	$41,144	$0.27	-3.55%	30391
800	Services	1977-1986	$132,285	$132,295	$6.12	-1.51%	30205
100	Agricultural services, forestry, fishing & other	1987-2000	$9,106	$9,108	$2.91	-8.98%	34424
200	Mining	1987-2000	$19,194	$19,197	$22.15	-54.35%	28985
300	Construction	1987-2000	$87,839	$87,835	$8.73	-7.39%	41675
400	Manufacturing	1987-2000	$268,234	$268,248	-$2.26	-10.99%	41703
500	Transportation and public utilities	1987-2000	$96,426	$96,423	$11.79	-4.64%	41850
610	Wholesale trade	1987-2000	$96,011	$96,023	$3.09	-6.93%	41106
620	Retail trade	1987-2000	$134,154	$134,158	-$1.17	-1.68%	43386
700	Finance, insurance, and real estate	1987-2000	$123,485	$123,508	$4.00	-5.14%	41747
800	Services	1987-2000	$398,884	$398,888	$23.26	-2.14%	42371

Table 2 continued. Accuracy of REIS income estimates (CA05 SIC) using "Average Proportion" method.

Line code	Description	Years	Percent of records in which the difference between the reported value and the estimate is within: ±$50	±$100	±$500	±$1000	±$5000	±$10000
100	Agricultural services, forestry, fishing & other	1969-1977	52.3%	70.0%	94.7%	98.1%	99.9%	100.0%
200	Mining	1969-1977	38.5%	52.7%	81.0%	88.5%	98.0%	99.4%
300	Construction	1969-1977	11.4%	20.7%	57.1%	72.3%	93.0%	96.9%
400	Manufacturing	1969-1977	10.3%	16.6%	42.7%	57.5%	86.5%	92.9%
500	Transportation and public utilities	1969-1977	20.0%	33.5%	73.1%	84.8%	96.4%	98.5%
610	Wholesale trade	1969-1977	20.2%	33.5%	72.2%	84.2%	96.5%	98.3%
620	Retail trade	1969-1977	15.8%	28.3%	70.0%	83.8%	96.5%	98.4%
700	Finance, insurance, and real estate	1969-1977	38.2%	55.8%	85.5%	91.6%	98.0%	99.0%
800	Services	1969-1977	14.3%	25.0%	63.0%	78.5%	95.1%	97.2%
100	Agricultural services, forestry, fishing & other	1977-1986	30.7%	47.3%	83.3%	91.4%	99.0%	99.7%
200	Mining	1977-1986	23.2%	34.6%	63.9%	74.6%	91.8%	96.0%
300	Construction	1977-1986	6.5%	12.0%	39.7%	55.5%	84.1%	91.3%
400	Manufacturing	1977-1986	4.9%	8.5%	25.4%	37.2%	70.4%	81.9%
500	Transportation and public utilities	1977-1986	8.6%	15.9%	48.6%	65.2%	89.6%	94.5%
610	Wholesale trade	1977-1986	11.0%	19.7%	54.0%	70.1%	92.2%	95.6%
620	Retail trade	1977-1986	7.0%	13.5%	46.3%	64.8%	90.8%	95.3%
700	Finance, insurance, and real estate	1977-1986	15.7%	27.3%	63.9%	77.1%	93.6%	96.5%
800	Services	1977-1986	5.5%	10.4%	36.3%	54.0%	85.8%	92.1%
100	Agricultural services, forestry, fishing & other	1987-2000	15.4%	27.1%	66.2%	80.5%	96.3%	98.6%
200	Mining	1987-2000	13.7%	23.1%	51.9%	65.0%	87.9%	93.6%
300	Construction	1987-2000	3.7%	7.0%	26.7%	41.2%	75.5%	85.8%
400	Manufacturing	1987-2000	2.5%	4.7%	16.8%	25.6%	57.2%	71.8%
500	Transportation and public utilities	1987-2000	4.3%	8.3%	30.8%	46.7%	79.9%	88.4%
610	Wholesale trade	1987-2000	5.6%	10.5%	36.2%	52.6%	82.9%	90.0%
620	Retail trade	1987-2000	4.1%	7.7%	29.8%	46.0%	79.8%	88.3%
700	Finance, insurance, and real estate	1987-2000	7.9%	14.6%	44.4%	59.4%	84.0%	89.7%
800	Services	1987-2000	2.7%	5.1%	20.9%	34.1%	69.6%	80.4%