

Linking firms with establishments in BLS microdata

An examination of the Employer Identification Numbers (EINs) of a sample of large firms that use multiple EINs reveals that only a small percentage of their establishments and employment can be identified with the one EIN that each such firm uses in its filing of Securities and Exchange Commission Form 10-K; approaches are suggested to overcome this limitation

Elizabeth Weber Handwerker
and
Lowell G. Mason

The Bureau of Labor Statistics (BLS, the Bureau) collects data from employers about their establishments. For certain applications, however, researchers inside and outside the Bureau need data on firms. For example, in an earlier *Monthly Labor Review* article, Elizabeth Handwerker, Mina Kim, and Lowell Mason attempted to find all of the establishments associated with the 500 largest multinational manufacturing firms identified in surveys conducted by the Bureau of Economic Analysis (BEA).¹ Other researchers have suggested merging BLS microdata with additional datasets containing information about firms.² This article (1) gives an overview of the complex relationship between firms, on the one hand, and their establishments and establishment identifiers, on the other, and (2) outlines the efforts involved in linking establishment data into firms.

The backbone of all employer microdata at the Bureau is the Quarterly Census of Employment and Wages (QCEW). Covering approximately 9 million establishments nationwide and 98 percent of U.S. employment, this dataset contains quarterly records of all U.S. business establishments subject to state Unemployment Insurance (UI) laws.³ The records of the QCEW include monthly

employment and quarterly total payroll data, based on the quarterly contribution reports employers submit to the state agencies responsible for administering UI programs. Each establishment in the QCEW is an economic unit, such as a farm, mine, factory, or store that produces goods or provides services. Establishments typically have a single physical location and are engaged in one type of economic activity.

In recent years, several researchers have expressed interest in merging corporate datasets compiled from firms' mandatory filings with the Securities and Exchange Commission (SEC) with QCEW data, using firms' federal Employer Identification Numbers (EINs) as the identifier for linking firm data to the establishment data of the QCEW. However, there is no simple way to use EINs to find, for a given firm, all of that firm's establishments in the confidential microdata of the QCEW. Although every establishment in the QCEW is associated with both a federal EIN and a state UI account number, businesses may use one EIN for the UI tax system and other, different EINs for other tax systems. Put another way, both EINs and UI account numbers define businesses for tax purposes, but a firm may have more than one EIN and more than one UI account number.

Elizabeth Weber Handwerker is a research economist in the Office of Employment and Unemployment Statistics, Bureau of Labor Statistics. Lowell G. Mason is an economist in the same office. Email: handwerker.elizabeth@bls.gov or mason.lowell@bls.gov.

Thus, firms may use one EIN in filings with the SEC and a different EIN (or set of EINs) in reporting to the UI system. Also, firms that span multiple states will have a different UI account in each state, and large, complex firms may use numerous EINs across many states.

The BLS Business Employment Dynamics program publishes estimates by firm size, based on QCEW data. These estimates, however, are calculated at the EIN level. In other words, there are no true firm identifiers, other than EINs and UI accounts, in the QCEW.

The body of this article begins by exploring the relationships among EINs, UI account numbers, and establishments in the QCEW. Next, the analysis goes on to examine a list of firms already matched with all their establishments in previous BLS efforts in order to show that the EINs which are readily available from firms' Form 10-K filings⁴ with the SEC link to only a subset of these firms' establishments. The analysis then discusses the methods and time required to link several case studies of sample firms to the full list of their establishments in BLS data. The article concludes with a brief synopsis of the material presented and sets forth a possible agenda for future research.

EINs, UI account numbers, and establishments in the QCEW

EINs are issued by the Internal Revenue Service to identify employers for tax purposes. As Joel Elvery, Lucia Foster, C.J. Krizan, and David Talan showed, most employers have only one EIN.⁵ In the fourth quarter of 2009, employers' reports to the UI system used 5.1 million EINs (although they were not necessarily the same EINs used in employers' reports to federal agencies, such as the SEC or the BEA). These same employers have 6.2 million accounts, covering 7.3 million establishments, in the UI system (with at least one account for each U.S. state and the District of Columbia).

Firms may use the same EIN in multiple states. However, as table 1 shows, 96 percent of EINs in the QCEW are associated with establishments in a single state. These EINs, each of which is associated with 1.1 establishments with a total of 11.3 employees, on average, contain 52.7 percent of all private sector employment covered in the QCEW. By contrast, only 0.4 percent of EINs are associated with establishments in 10 or more states, but these EINs are associated with an average of 52.8 establishments each, with an average total employment of 1,690.0 employees, representing 28.9 percent of all covered employment in the QCEW.

Similarly, table 2 shows that 94.9 percent of EINs are associated with a single establishment, but these EINs account for just 42.2 percent of private sector employment

in the QCEW. Meanwhile, the 0.7 percent of EINs that are associated with 10 or more establishments have an average of 50.8 establishments each, and these EINs make up 40.3 percent of all private sector employment.

Additional information on the distribution of UI accounts across states and the distribution of establishments within UI accounts is shown in tables 3 and 4, respectively. Table 3 shows that 95.9 percent of EINs are associated with a single UI account, but these EINs account for 52.2 percent of all private sector employment included in the QCEW. Meanwhile, the 0.4 percent of EINs that are associated with 10 or more UI accounts constitute 29.3 percent of all private sector employment. Recall from table 1 that nearly all of the 4,933,965 EINs in the QCEW that are associated with establishments in a single state are associated with a single UI account. (Only 3,695, or 0.1 percent, are associated with more than one UI account.) Table 4 shows that 98.1 percent of UI accounts are associated with a single establishment and that 82.6 percent of establishments hold single-establishment UI accounts. These accounts are associated with 61.5 percent of all employment covered in the QCEW. However, the 0.4 percent of UI accounts that are associated with at least 10 establishments are associated with 12.6 percent of the establishments, and 23.2 percent of the employees, in the QCEW. These UI accounts are each associated with an average of 34 establishments, which tend to be larger than the establishments holding single-establishment UI accounts.

Chart 1 shows the number of EINs, UI accounts, and private establishments in the QCEW from the first quarter of 1991 through the fourth quarter of 2011. The chart reveals that all of these measures are increasing over time, with faster growth in the number of establishments than UI accounts and in the number of UI accounts than EINs.

During the period shown, the quarterly growth rates of EINs, UI accounts, and establishments are seen to be roughly correlated. Overall, the "complexity" of companies in terms of the number of establishments per UI account and per EIN increased from 1991 to 2011. For researchers who are searching for establishments associated with particular companies, this means that the average number of establishments that can be linked to each EIN has been increasing. However, that fact does not help researchers who are searching for all of the EINs associated with large firms.

Establishments, employment, and EINs in firms' public filings

Publicly held firms are required to report to the SEC. The information they report (particularly in Form 10-K) is of interest to many researchers and is compiled into com-

Table 1. Distribution of Employer Identification Numbers (EINs) across states, fourth quarter, 2009

Number of states	Number of EINs (percentage of total)	Number of establishments (percentage of total)	Average number of establishments per EIN (standard deviation)	Sum of employment for all EINs in category (percentage of total)	Average employment per EIN (standard deviation)
Total	5,141,516 (100.0)	7,336,839 (100.0)	1.4 (124.8)	106,104,761 (100.0)	20.6 (770.4)
1	4,933,965 (96.0)	5,580,384 (76.1)	1.1 (126.5)	55,947,126 (52.7)	11.3 (156.7)
2	114,970 (2.2)	309,084 (4.2)	2.7 (7.1)	6,946,347 (6.5)	60.4 (398.0)
3	32,375 (.6)	143,131 (2.0)	4.4 (25.9)	3,448,774 (3.3)	106.5 (477.1)
4	15,656 (.3)	93,710 (1.3)	6.0 (22.5)	2,459,057 (2.3)	157.1 (1,262.8)
5	9,484 (.2)	70,486 (1.0)	7.4 (19.3)	1,807,086 (1.7)	190.5 (746.3)
6	6,314 (.1)	55,462 (.8)	8.8 (25.2)	1,509,262 (1.4)	239.0 (1,117.3)
7	4,572 (.1)	47,360 (.6)	10.4 (26.9)	1,161,456 (1.1)	254.0 (750.4)
8	3,377 (.1)	39,438 (.5)	11.7 (28.7)	1,011,136 (1.0)	299.4 (1,165.4)
9	2,643 (.1)	39,016 (.5)	14.8 (49.9)	1,123,213 (1.1)	425.0 (1,799.5)
10 or more	18,160 (.4)	958,768 (13.1)	52.8 (232.7)	30,691,304 (28.9)	1,690.0 (12,412.9)
25 or more	4,574 (.1)	619,973 (8.5)	135.5 (435.7)	21,155,404 (19.9)	4,625.1 (24,195.2)

SOURCE: U.S. Bureau of Labor Statistics.

mercial databases used by many researchers. Each firm's Form 10-K report includes one EIN, which is included in those databases. Several researchers have proposed research projects that would merge a commercial database of firm information with QCEW data, using only the single EIN per firm listed in the commercial database. However, firms may use many different EINs for different purposes, and many firms use multiple EINs in reporting unemployment insurance taxes. The EIN that a firm reports to the SEC in Form 10-K may be one of many EINs associated with establishments of the same firm in the QCEW or may even be an EIN never used in the QCEW. The analysis that follows uses only the EINs that these firms list in their Forms 10-K to examine the percentage of establishments and the percentage of employment that

can be linked to a list of large firms.

The comparison presented of the total number of establishments and employees with the number that can be linked to the one EIN listed in each firm's Form 10-K is based on a list of firms whose EINs BLS analysts believe that they know. The list was developed at the Bureau to avoid sampling only one part of a large employer in surveys. Forty-three large publicly held firms appearing in this list are examined. (The full list contains information on more than two hundred firms; this article uses all of the firms from the list that were part of the Dow Jones Industrial Average or the Standard and Poor's (S&P) 500® Index, as well as a random sample of other firms on the list that were included in the Russell 2000 Index.) The following tabulation shows the percentage of estab-

Table 2. Distribution of establishments within Employer Identification Numbers (EINs), fourth quarter, 2009

Number of states	Number of EINs (percentage of total)	Sum of establishments for each category of establishments (percentage of total)	Average number of establishments per EIN (standard deviation)	Sum of employment for each category of establishments (percentage of total)	Average employment per UI account (standard deviation)
Total	5,141,516 (100.0)	7,336,839 (100.0)	1.4 (124.8)	106,104,761 (100.0)	20.6 (770.4)
1	4,877,459 (94.9)	4,877,459 (66.5)	1.0 (.0)	44,794,423 (42.2)	9.2 (44.9)
2	125,147 (2.4)	250,294 (3.4)	2.0 (.0)	5,460,496 (5.1)	43.6 (157.9)
3	40,479 (.8)	121,437 (1.7)	3.0 (.0)	3,278,666 (3.1)	81.0 (299.4)
4	22,022 (.4)	88,088 (1.2)	4.0 (.0)	2,327,721 (2.2)	105.7 (328.6)
5	14,546 (.3)	72,730 (1.0)	5.0 (.0)	1,970,604 (1.9)	135.5 (414.4)
6	10,364 (.2)	62,184 (.8)	6.0 (.0)	1,695,390 (1.6)	163.6 (487.4)
7	7,340 (.1)	51,380 (.7)	7.0 (.0)	1,396,653 (1.3)	190.3 (625.5)
8	5,689 (.1)	45,512 (.6)	8.0 (.0)	1,273,174 (1.2)	223.8 (1,012.5)
9	4,490 (.1)	40,410 (.6)	9.0 (.0)	1,123,669 (1.1)	250.3 (638.6)
10 or more	33,980 (.7)	1,727,345 (23.5)	50.8 (1,534.5)	42,783,965 (40.3)	1,259.1 (9,338.4)
25 or more	10,500 (.2)	1,382,658 (18.8)	131.7 (2,758.8)	33,292,362 (31.4)	3,170.7 (16,592.4)

SOURCE: U.S. Bureau of Labor Statistics.

lishments and the percentage of employment that can be identified by the single EIN listed in firms' Form 10-K filings, by category:

Category	Number of firms examined	Percentage of establishments	Percentage of employment	<i>Firms for which the single EIN listed in Form 10-K is used by establishments with North American Industrial Classification (NAICS) code 551</i>
Dow Jones Industrial Average	14	4.1	0.8	9
S&P 500 (excluding Dow Jones)	14	22.3	2.1	13
Russell 2000	15	42.2	3.0	7

The categories used in the preceding tabulation are the indexes in which the firms are listed. Many of the largest publicly held companies in the United States are included in the Dow Jones Index. The S&P 500 Index includes the 500 largest publicly held companies in the nation (chosen by a committee that examines various measures of firm size), while the Russell 2000 Index excludes the largest 1,000 companies and includes companies ranked 1,001 to 3,000 in size (by market capitalization). Thus, the categories used in the tabulation are a rough indication of the size of the companies examined in this article.

The percentages of establishments and employment that can be identified with the single EIN used in firms' Form 10-K filings are least for the firms listed in the Dow Jones Industrial Average (the largest and most complex

Table 3. Distribution of Unemployment Insurance (UI) accounts within states and Employer Identification Numbers (EINs), fourth quarter, 2009

Number of UI accounts	Number of EINs (percentage of total)	Number of establishments (percentage of total)	Average number of establishments per UI account (standard deviation)	Sum of employment for employers in each category (percentage of total)	Average employment per UI account (standard deviation)
Total	5,141,516 (100.0)	7,336,839 (100.0)	1.4 (124.8)	106,104,761 (100.0)	20.6 (770.4)
1	4,930,270 (95.9)	5,253,433 (71.6)	1.1 (2.3)	55,354,699 (52.2)	11.2 (89.7)
2	118,214 (2.3)	315,574 (4.3)	2.7 (7.1)	7,051,820 (6.6)	59.7 (405.0)
3	32,586 (.6)	143,332 (2.0)	4.4 (25.8)	3,474,382 (3.3)	106.6 (477.2)
4	15,718 (.3)	93,746 (1.3)	6.0 (22.5)	2,475,125 (2.3)	157.5 (1,262.0)
5	9,502 (.2)	70,833 (1.0)	7.5 (19.3)	1,831,208 (1.7)	192.7 (752.1)
6	6,347 (.1)	54,106 (.7)	8.5 (20.6)	1,501,898 (1.4)	236.6 (1,107.1)
7	4,589 (.1)	47,430 (.6)	10.3 (26.9)	1,158,410 (1.1)	252.4 (733.1)
8	3,392 (.1)	39,496 (.5)	11.6 (28.5)	1,047,219 (1.0)	308.7 (1,230.3)
9	2,645 (.1)	38,821 (.5)	14.7 (49.8)	1,106,241 (1.0)	418.2 (1,785.6)
10 or more	18,253 (.4)	1,280,068 (17.4)	70.1 (2,092.6)	31,103,759 (29.3)	1,704.0 (12,554.4)
25 or more	4,648 (.1)	948,075 (12.9)	204.0 (4,142.4)	21,693,672 (20.4)	4,667.3 (24,353.4)

SOURCE: U.S. Bureau of Labor Statistics.

firms) and greatest for the firms listed in the Russell 2000 Index (the smallest examined in this article). This situation suggests that the larger a publicly held firm, the smaller are the percentages of its establishments and employment that can be identified with the single EIN listed in the firm's Form 10-K filings. Still, even for the smallest of the publicly held firms on the list (those listed in the Russell 2000 Index), less than half of all establishments and a very small percentage of employment can be linked directly to the QCEW by using only the EINs listed in the firms' Form 10-K filings. Those EINs frequently can be linked with establishments of firms classified into NAICS code 551, "Management of Companies." The following tabulation gives the actual number of EINs used by these firms, as well as the number of states in which

the firms operate:

Category	Number of firms examined	Mean number of EINs per firm (standard deviation)	Mean number of states with nonzero employment per firm (standard deviation)	Mean number of states per EIN (standard deviation)
Dow Jones Industrial Average.....	14	29.5 (39.8)	48.9 (4.5)	13.4 (15.8)
S&P 500 (excluding Dow Jones) ...	14	231.5 (869.7)	36.5 (18.8)	2.3 (6.7)
Russell 2000 ...	15	5.0 (4.4)	34.2 (18.5)	16.2 (18.3)

Table 4. Distribution of establishments within Unemployment Insurance (UI) accounts, fourth quarter, 2009

Number of establishments	Number of UI accounts (percentage of total)	Sum of establishments for each category of UI accounts (percentage of total)	Average number of establishments per UI account (standard deviation)	Sum of employment for each category of UI accounts (percentage of total)	Average employment per UI account (standard deviation)
Total	6,177,029 (100.0)	7,336,839 (100.0)	1.2 (5.2)	106,104,761 (100.0)	17.2 (211.7)
1	6,060,855 (98.1)	6,060,855 (82.6)	1.0 (.0)	65,270,467 (61.5)	10.8 (63.4)
2	28,005 (.5)	56,010 (.8)	2.0 (.0)	3,699,202 (3.5)	132.1 (432.8)
3	18,756 (.3)	56,268 (.8)	3.0 (.0)	3,049,493 (2.9)	162.6 (617.6)
4	12,878 (.2)	51,512 (.7)	4.0 (.0)	2,358,709 (2.2)	183.2 (608.9)
5	9,656 (.2)	48,280 (.7)	5.0 (.0)	1,980,755 (1.9)	205.1 (551.1)
6	7,343 (.1)	44,058 (.6)	6.0 (.0)	1,710,925 (1.6)	233.0 (574.3)
7	5,064 (.1)	35,448 (.5)	7.0 (0.0)	1,291,310 (1.2)	255.0 (617.1)
8	4,025 (.1)	32,200 (.4)	8.0 (.0)	1,157,933 (1.1)	287.7 (714.1)
9	3,256 (.1)	29,304 (.4)	9.0 (.0)	967,459 (.9)	297.1 (632.5)
10 or more	27,191 (.4)	922,904 (12.6)	33.9 (71.3)	24,618,508 (23.2)	905.4 (2,714.2)
25 or more	9,211 (.1)	657,623 (9.0)	71.4 (113.3)	16,264,898 (15.3)	1,765.8 (4,297.8)

SOURCE: U.S. Bureau of Labor Statistics.

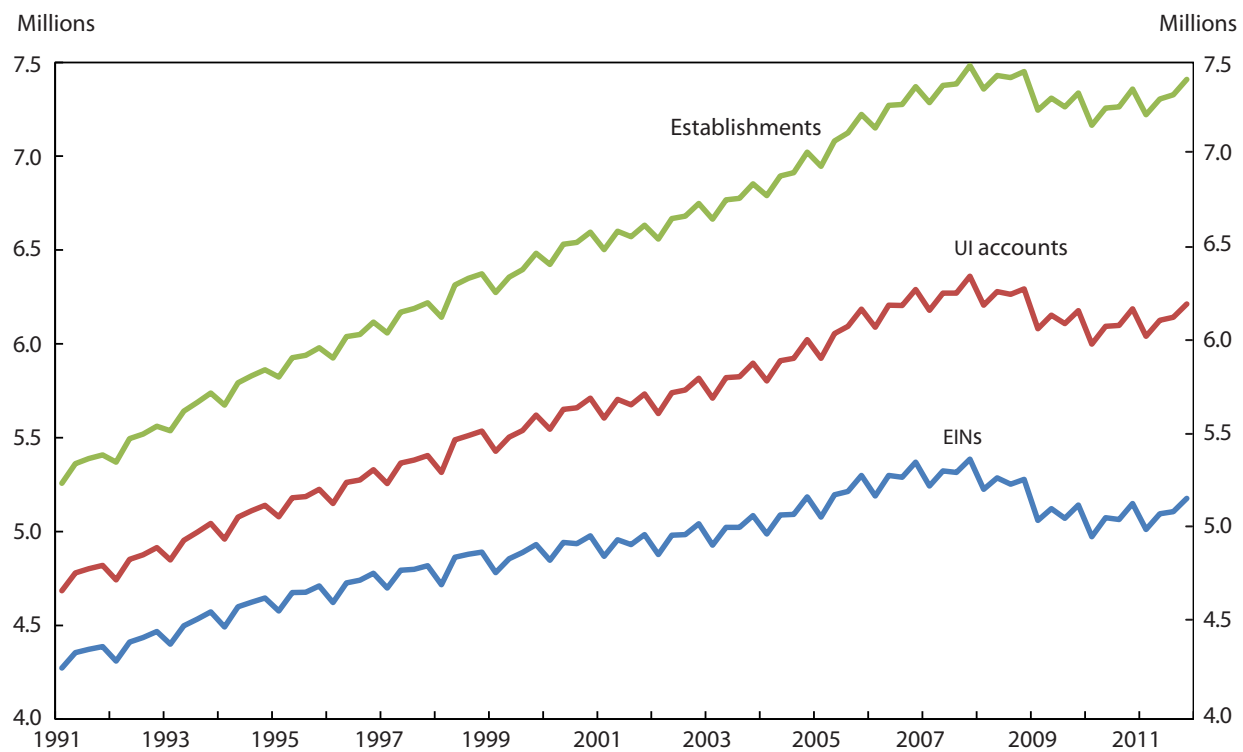
Case studies

Finding all the establishments (in practical terms, finding all the EINs associated with the establishments) for a firm appearing in BLS data is important to researchers who want to link firm-level data with BLS establishment-level microdata. This section presents four case studies of the efforts involved in such linking. Examined are one firm listed in the Dow Jones Industrial Average, one firm listed in the S&P 500 Index, one firm listed in the Russell 2000 Index, and one firm that is privately held (and thus would not need to file Form 10-K). These firms were not chosen completely at random; rather, they were selected because information on their total employment was available (in some cases, from the firm's website or Form 10-K

filing). For each firm, a certain percentage of establishments and the percentage of employment can be found by searching for the firm name in the QCEW for the fourth quarter of 2009. Greater percentages of establishments and employment can be found through more rigorous matching efforts that use the names of all subsidiaries and all addresses of establishments of the firms listed in Form 10-K reports for 2009 and on firms' websites. However, more time is required to find these additional names and addresses.

Searching for establishments by firm name has advantages and disadvantages for researchers. The QCEW contains both legal and trade names for each establishment, and these names can be used in computer searches. However, many of the names listed in the QCEW are

Chart 1. Number of Employer Identification Numbers (EINs), Unemployment Insurance (UI) accounts, and establishments, first quarter, 1991, through fourth quarter, 2011



SOURCE: U.S. Bureau of Labor Statistics.

older names of company plants or subsidiaries. Moreover, few names are unique, so, in addition to matching the establishments found with the firm in question, computer searches for establishments by firm name may incorrectly match the name with the establishments of hundreds or thousands of other firms. Another drawback is that the QCEW includes only the most recent version of names and addresses of establishments, complicating name searches for establishments that operated during earlier periods.

The following tabulation, for the fourth quarter of 2009, shows the percentage of establishments and total employment matched when just the firm name was used, for each of the sample firms involved in the four case studies:⁶

Sample firm studied	Establishments		Employment	
	Percentage correctly matched	Percentage incorrectly matched	Percentage correctly matched	Percentage incorrectly matched
1: listed in Dow Jones.....	80.5	170.0	95.8	38.8
2: listed in S&P 500.....	4.8	0.0	2.4	0.0
3: listed in Russell 2000.....	99.0	0.0	97.6	0.0

4: privately owned..... 0.8 0.0 52.8 0.0

In only two of the four case studies—the Dow Jones and Russell 2000 listings—was a large percentage of establishments and employment correctly identified. However, the Dow Jones listing also had a large percentage of incorrectly identified establishments and total employment. Reviewing the resulting matches in this case study revealed that a number of establishments were acquired by the firm after the fourth quarter of 2009 and that these establishments were incorrectly matched. (The QCEW name and address files are continuously updated, and versions corresponding to past dates are not available.) The remaining two case studies identified much smaller percentages of establishments and employment. A review of the establishments that were not identified indicated that the unmatched establishments' names listed in the QCEW were those of the associated firms' subsidiaries and not the firms themselves.

As the following tabulation shows, better results can be obtained by using the names of firms' subsidiaries as well as the addresses of the firms' establishments:

<i>Sample firm studied</i>	<i>Establishments</i>		<i>Employment</i>	
	<i>Percentage correctly matched</i>	<i>Percentage incorrectly matched</i>	<i>Percentage correctly matched</i>	<i>Percentage incorrectly matched</i>
1: Listed in Dow Jones.....	99.7	0.3	100.0	0.0
2: listed in S&P 500	100.0	.0	100.0	.0
3: listed in Russell 2000	100.0	.0	100.0	.0
4: privately owned.....	99.8	.0	99.9	.0

These names and addresses are culled from the firms' websites and from their Form 10-K filings (for each firm that is publicly listed), but this manual process is time consuming. To aid the process, the search may be expanded to include establishments in the QCEW with names that do not exactly match those of the associated firms and their subsidiaries, but rather match only parts of the names. This approach increases the number of possible matches, both correct and incorrect. Thus, the matches are reviewed manually and compared against addresses found on the firm's Form 10-K listings and websites, and then the incorrect matches are removed. As the following tabulation shows, this additional manual step adds more time to the matching process (but it is much better than simple searches by name or single EIN):

<i>Sample firm studied</i>	<i>Minutes spent—</i>			
	<i>Reviewing firms' Form 10-K listings</i>	<i>Reviewing firms' websites</i>	<i>Searching the QCEW by subsidiaries' names and addresses</i>	<i>Total time taken (minutes)</i>
1: listed in Dow Jones.....	10	21	27	58
2: listed in S&P 500.....	26	39	21	86
3: listed in Russell 2000....	13	10	18	41
4: privately owned.....	0	44	25	69
Average minutes spent.....	12	29	23	64

Still, these efforts do not find every correct match or remove every incorrect match. Fortunately, for the four case studies presented, there is additional information about

the true matches, and that information can be used to evaluate the matching efforts. (Note, however, that, for most firms, such information is not available.)

In each of the four case studies, all of the firm's establishments were found. Not every matching attempt, however, is successful. For example, Handwerker, Kim, and Mason attempted to find all of the establishments in the QCEW for the largest 500 multinational manufacturers in the United States.⁷ Using every resource currently available at the Bureau, they were able to find establishments that matched employment within 20 percent of total employment reported to BEA for only 454 of the firms examined.

RESEARCHERS SOMETIMES NEED TO FIND all of the establishments associated with a single employer in BLS data. With most employers, this task for the researcher is straightforward. As shown in tables 1 through 4 and by Elvery and colleagues,⁸ the vast majority of employers are small, with EINs in only one state and with a single UI account and a single establishment. However, the large companies that frequently are of interest to researchers often use multiple EINs in reporting their employment to the UI system (the source of QCEW data), and there is no straightforward way to find all of the EINs and establishments associated with a particular firm.

This article has examined a sample of large firms and found that only a small percentage of these firms' establishments and employment can be identified by using the one EIN that each firm reports in its Form 10-K filings with the SEC. To determine the effort needed to identify the EINs (and thus establishments) of all firms, case studies of sample firms were undertaken. With information culled from firms' Form 10-K filings, almost all of these sample firms' establishments and total employment were able to be identified, with about an hour's work of searching and verifying per firm. Still, as noted by Handwerker, Kim, and Mason, such efforts are not always successful.⁹

Under a new agreement with the Census Bureau, the Bureau of Labor Statistics will soon receive Census Bureau data from that agency's Company Organization Survey on EINs that make up large companies. Of future interest will be whether this newly shared data substantially reduces the effort required to find all the establishments of large companies in BLS data. □

Notes

¹ Elizabeth Weber Handwerker, Mina M. Kim, and Lowell Mason, "Domestic employment in U.S.-based multinational companies," *Monthly Labor Review*, October 2011, pp. 3–15.

² On a limited basis, the Bureau allows eligible researchers to access confidential data for purposes of conducting valid statistical analyses. (For more information, see "BLS information: researcher access to con-

confidential data files at the BLS” (U.S. Bureau of Labor Statistics, Aug. 22, 2011), <http://www.bls.gov/bls/blsresda.htm>.)

³ Employment data exclude members of the Armed Forces, the self-employed, proprietors, domestic workers, unpaid family workers, and railroad workers covered by the railroad Unemployment Insurance system.

⁴ SEC Form 10-K “provides a comprehensive overview of the company’s business and financial condition and includes audited financial statements.” (See “Form 10-K” (Securities and Exchange Commission, June 26, 2009), <http://www.sec.gov/answers/form10k.htm>.)

⁵ Joel Elvery, Lucia Foster, C.J. Krizan, and David Talan, “Prelimi-

nary micro data results from the Business List Comparison Project,” *Proceedings of the 2006 American Statistical Association Annual Meeting* (Alexandria, VA, American Statistical Association, 2006).

⁶ To verify in all four cases that the establishments that were found through the matching process used were the correct establishments, the names of establishments were examined and the EINs that were found by matching against the (highly incomplete) BLS listings of EINs for employers with multiple EINs were checked.

⁷ Handwerker, Kim, and Mason, “Domestic employment.”

⁸ Elvery, Foster, Krizan, and Talan, “Preliminary Micro Data Results.”

⁹ Handwerker, Kim, and Mason, “Domestic employment.”