Linking Foreign Direct Investment Data (FDI) to the BLS Business Register: Developing Employment Measures Related to Foreign Direct Investment

David M. Talan

Lowell G. Mason

Richard L. Clayton

U.S. Bureau of Labor Statistics

Motivation

One data gap that Bureau of Labor Statistics (BLS) users mention often is the need to measure the employment effects from inward foreign direct investment (FDI). Since building new surveys is expensive and can take several years to implement, we propose to create these measures by linking existing Bureau of Economic Analysis (BEA) reports on foreign direct investment with BLS business register data.

The BLS uses its Quarterly Census of Employment and Wages (QCEW) Business Register to publish measures of employment and wage levels, employment dynamics and measures of entrepreneurship which are released quarterly. The QCEW also serves as the frame for BLS surveys, such as the Occupational Employment Survey (OES). The QCEW is a comprehensive source of data on U.S. establishments. At the establishment level, total employment and compensation data are collected, as well as the industrial classification and geographic location of each establishment.

BEA collects information on foreign direct investment (FDI) through a combination of census years and subsets in intervening years. BEA conducts these programs to measure U.S. assets overseas and foreign ownership of U.S. assets. BEA produces statistics on FDI in the U.S. and US direct investment abroad that are essential to the compilation of the U.S. economic accounts and for the analysis of multinational enterprises (MNEs). The statistics are obtained from mandatory surveys. BEA collects data at the enterprise level.

This paper outlines the conceptual framework for developing indicators of employment and wages related to FDI by linking micro-data from BEA with the QCEW business register. There are many reasons to do so. First, this approach leverages existing data, relieving response burden. Further, BLS establishment data augment the enterprise-level BEA FDI data, allowing a detailed look at the distribution of establishments (across industry or geographic location, for instance) that comprise foreign-owned enterprises, as well as the ability to compare those to domestically-owned establishments. Lastly, as the QCEW is the frame for the OES, it is also possible to track the distribution of occupations within establishments that are foreign-owned.

Objectives

First, we briefly describe the inputs used to create the BLS business register. Then we describe how we construct our business register in section II. In section III, we describe the BEA dataset on foreign direct investment. Section IV describes our matching process. The quality of the matches is then described in section V. Finally, we conclude with opportunities for future work.

Section I: Inputs of the QCEW Business Register

The QCEW Business Register is a list of active employer business establishments in the United States, Washington DC, Puerto Rico, and the Virgin Islands. Its principal sources of information are the mandatory quarterly reports filed by all employers covered by the Unemployment Insurance system of the fifty States, the District of Columbia, Puerto Rico and the Virgin Islands. Employers report to their

State Workforce Agencies (SWA's) in compliance with State Unemployment Insurance (UI) laws, and for Federal civilian workers, in compliance with the Unemployment Compensation for Federal Employee (UCFE) program. Each quarter, business and government employers report monthly employment and quarterly wages.

The QCEW Business Register contains data collected from four forms required to meet program needs. Two of these are state Unemployment Insurance based forms: the first is the Status Determination Form (SDF) and the second is the Quarterly Contribution Report (QCR). In addition, two BLS supplemental collections were developed to enhance the administrative data to meet the needs of the QCEW program. The first of the BLS forms is the Annual Refiling Survey (ARS) and the second is the Multiple Worksite Report (MWR). The purposes of these forms will be explained in more detail later in this section.

The Council of Professional Associations on Federal Statistics (COPAFS) surveyed their members in 2015 seeking to profile the US statistical system's administrative databases. The BLS QCEW program, and thus the BLS business register, was listed as the number one most important administrative datasets in the US statistical system. The FDI dataset produced by BEA was listed as the most important international dataset. Thus, this project matches two very important datasets to produce new statistical measures.

Maintenance Workload

To gain some perspective of the size and workload associated with maintaining the QCEW Business Register, there are currently 9.4 million worksites representing 139 million employees reported as of the fourth quarter of 2014.

Approximately 7.5 million worksites were reported as being single establishment employers, whose employment and wages data were collected from the state QCRs. The employment and wages data and other business identification information for the remaining 1.9 million worksites were collected from the MWRs.

There are approximately 7 million legal entities with only about 136,000 of those providing the MWR data. The term "legal entity" is used here since many large employers have accounts in more than one state and are thus counted more than once. The MWR employers represent about 1.4% of total employers, but they constitute 17% of the total number of worksites and a stunning 41% of the Nation's employment. Thus, without the MWR, we could not accurately measure data about establishment births and deaths or accurately allocate employment and wages by industry and geography.

The Multiple Worksite Report is designed to collect information showing the distribution of the employment and wages of business establishments by industry and geographic area. Information on the MWR form is used to more accurately classify employment and wage data of multiple establishment employers by industry and by location within a State. By collecting and storing employment and wage data by worksite, States can disaggregate these data below the county level for more extensive and detailed analysis of business and economic conditions within their State, including local and regional employment totals. These data are used to ensure an equitable distribution of Federal funds through

grant programs that use county economic indicators as a basis for allocations. No other sources are available to obtain this information.

Status Determination Form (SDF)

All new employers which become subject to UI coverage are required to file a Status Determination Form with the UI unit of the State Employment Security Agency (SESA). This form (or screen for respondents who report via the web) is used to determine an employer's tax liability under the State's UI laws and to collect administrative information. The form includes basic identification information, including business name, mailing and physical location address, type of organization, Federal Employer Identification Number (EIN), and more. The Status Determination Form also requests information to identify the establishment's industrial activity (NAICS) code, county code (or township in the New England area), and the ownership code (private sector, or Federal, State or local government). These codes are assigned by the SESA staff. The assignment of the industry code is based on the establishment's primary economic activity, which is determined by its principal product or group of products produced or distributed, or services rendered. If there is insufficient information on the Status Determination Form, the employer will either be contacted by telephone or mailed an Industry Verification Form to obtain the necessary information. In a sense, it is the State's equivalent of Social Security's Form SS-4, Application for a Federal EIN.

The QCEW program contains an EIN because of the Federal Unemployment Tax Act (FUTA). The Federal Unemployment Tax Act, along with state unemployment systems, provides for payments of unemployment compensation to workers who have lost their jobs. Most employers pay both a federal and a state unemployment tax. FUTA law provides for a substantial offset for federal taxes for employers that pay into State UI taxes. Thus, employers have a tremendous incentive to report the same EINs to both State and federal systems. Because the employer pays the FUTA tax; the EIN can be matched between the IRS and the state to verify that these payments have been made. The EIN is the key variable for matching records in this project.

Quarterly Contribution Report (QCR)

All liable employers are required to file a QCR with the SESAs for their UI accounts. These reports, like the Status Determination Forms, are administered by the UI program. All of the QCR forms request employment values for each month of the quarter and total wages, taxable wages, and UI taxes due for the quarter. This information and the taxes that are due are necessary for the operation of the UI tax system, but they are also an important source of statistical data. Employers are asked to report, among other items, the total number of covered workers (full and part-time) who earned wages (subject to UI taxes) during the pay period(s) which includes the 12th of each month in the quarter and the total payroll for the quarter. This is the official OMB definition of employment. This report is mandatory for employers with a single location as well as employers with multiple locations in the State. The latter group of employers reports a summary of these data for all of their establishments covered under the same State UI account on the QCR.

Multiple Worksite Report (MWR)

Business enterprises with more than one establishment in a State under a single UI account file a Multiple Worksite Report (MWR) so that data for each of its establishments is reported separately. The MWR is mandatory in 26 states and compliance in the voluntary states is also very high. The EIN provides linkages among establishments of the same business enterprise across States. Also collected are the physical location address of each worksite, a worksite description (normally a store or unit number or other information meaningful to the employer), and various other business identification information. This more comprehensive disaggregation of multi-establishment accounts is available in the QCEW Business Register which is almost entirely at the establishment level and thus provides more accurate industrial and geographic information for all establishments.

Since UI coverage is virtually 100 percent in all industries, the employment and wages from this program represent a virtual census of the Nation's employment and wages each quarter.

Data collection procedures for multi-establishment employers differ from those for single units. New worksites are added to the MWR by employers and closed sites are noted. Thus, new births and deaths and their associated employment are identified each quarter in a timely manner. This rapid identification of births and deaths improves the sample frame. Thus, the MWR captures business births and deaths for these multi-establishment employers on an on-going quarterly basis.

Annual Refiling Survey (ARS)

The purpose of the BLS Annual Refiling Survey is to review and, if necessary, update the classification codes (industrial, geographical, and ownership) currently assigned to the establishments stored on the QCEW Business Register. The survey covers approximately one-third of the establishments in the QCEW Business Register which are reviewed annually. Due to budget limitations, establishments with zero to three employees are excluded. The establishments are selected at random based on the Federal EIN range. This selection process ensures that the industrial distribution of the survey respondents is proportional to the establishments in the economy. In other words, no industrial sector is specifically targeted in any one year. In addition to the industrial classification review, the respondent is also requested to review and update information for the business establishment's 1) physical location address; 2) mailing address; and, 3) county, township, island, or parish in which the establishment is located.

Editing Procedures

Micro data collected on the QCR, MWR, and RFEW are edited by the SESA staff and corrected, as necessary. The micro data, including imputed values, are then aggregated to the appropriate macro-level cells. The SESA unit then edits and updates these macro records. There are 150 separate edits designed to detect a wide range of invalid and inconsistent values. These edits have been refined and enhanced over time reflecting the many years of data editing experience of state, regional and national office staff. Both the micro and macro edits include checks for invalid and inconsistent data as well as checks for large or unusual employment changes -- including over the month, over the quarter and over

the year -- and similar wages fluctuations between and within quarters. These edits and continuous scrutiny insure that QCEW employment and wages are of the highest accuracy.

Section II: Construction of the QCEW Business Register

In order to assure accurate linkages between new and existing establishments, there are two components of QCEW Business Register matching process: automated matching and analyst matching. The computer automated matching links approximately 97 percent of all establishments each quarter through a process which links State Employment Security Agencies' identification numbers (SESA-ID). Less than one percent of records are linked by statistical weighted matching or analyst matching. Every time a link is made, an identifier is assigned to the link to distinguish which type of match was made such as a breakout or consolidation.

Identifying Births and Deaths

The QCEW Business Register captures approximately 200,000 new establishment births and deaths every quarter. By collecting timely data on births and deaths, a number of BLS survey programs are able to include these establishments in their sample, thus making them more representative of the universe. Since these new births and deaths are available immediately, this reduces survey costs and leads to higher initial response rates.

SESA ID Matching

Ninety seven percent of all BLS establishment records are linked through a process by State Employment Security Agencies' identification numbers (SESA-ID). The SESA-ID is the establishment's unique identifier that the State Employment Security Agencies transmit to BLS, and the QCEW Business Register initially matches all units with identical SESA IDs; these matches are continuous establishments from the previous quarter.

Probability Weighted Matching

Probability weighted matching is a process by which related units are linked based on the similarity of various characteristics between two records. Pairs of records having enough in common are scored to exceed a specific point value and are then identified as valid matches. This point value is called a cutoff weight, and do not vary upon State, industry, or other characteristics, and create a consistent measure across the nation.

Analyst Manual Matching

As a final quality control measure of QCEW Business Register data, each quarter an analyst reviews data that contain records that were not matched by the automated process. The reasons for this additional review are that data elements may be miscoded, whereby inaccurate information is placed in system identifiers that the automated processes targets and therefore the system cannot make a linkage. Another reason for the supplementary analyst review is that certain records should not be linked, even though the system identifies a weighted match, and should be added to the QCEW Business Register as

an establishment birth or removed as an establishment death. The analyst manual review is essential to maintaining proper linkages and to preserve the high data quality produced by the QCEW Business Register.

Section III: BEA Inward FDI dataset

The BEA-provided data on US affiliates of foreign multinational enterprises comes from the 2012 Benchmark Survey of Inward Foreign Direct Investment in the United States[1]. The benchmark survey (every 5 years) is a mandatory reporting of all U.S. business enterprises (from here on referred to as affiliates) in which there is foreign direct investment by an entity (e.g. an individual, corporation, estate, trust, or government) that has at least a 10 percent controlling interest in the affiliate. The reports captures a comprehensive set of data, including:

- Balance sheet details, such as total assets, liabilities, and owner's equity
- Value added
- Employment and employee compensation
- Sales and net income
- Expenditures and trade in goods
- Expenditures in research and development

For the purposes of linking the BEA FDI dataset with QCEW establishment data, affiliate employment totals provide a target to measure how successfully establishments were matched to their parent affiliates. BEA employment represents the number of full and part-time employees on the affiliate's payroll at the end of the fiscal year. The BEA definition of employment is similar but not identical to the BLS. We use the same reference month when comparing BEA affiliate employment and the sum of employment for all matching QCEW establishments.

For this matching project, BLS is focusing on the US affiliates of foreign-owned multinationals in 2012 for which there is non-zero employment. These are the group of affiliates for which establishment matches are attempted. The average employment of these affiliates is 688.50 (standard deviation is 4,627.01). There is considerable variation among average affiliate employment (Table 1). 35.82% of all affiliates are quite small, employing less than 10 employees; they account for less than 0.18% of total employment for the affiliates with positive employment. Affiliates with 2,500 or more employees, on the other hand, account for 75.82% of total employment.

---

[1] A detailed description of the data is available in the BEA publication "Foreign Direct Investment in the United States (FDIUS): Preliminary 2012 Data," found at http://www.bea.gov/international/fdius2012_preliminary.htm.

Table 1: Size distribution of US affiliates of foreign-owned enterprises

Total number of employees	Percent of affiliates	Percent of employment
1 – 9	35.82%	0.18%
10 – 19	9.87%	0.20%
20 – 99	22.78%	1.61%
100 – 249	10.33%	2.40%
250 – 999	12.32%	9.20%
1000 – 2499	4.50%	10.61%
2500+	4.38%	75.82%

Country Group of UBO of Affiliate

Tables 2 shows the 3 top groups of countries of UBO for the affiliates with positive employment that are eligible to be matched both in terms of the number of affiliates as well as total employment. As shown in Table 2, Europe accounts for the largest percentage of foreign-owned US affiliates, both in terms of number of affiliates (44.28%) and total employment (63.48%). On average, US affiliates whose ultimate beneficial ownership are European enterprises are far larger than those of other country groups 987.21 employees per affiliate as compared to 688.50 for all affiliates). Asia and Pacific owned affiliates account for the next highest percentages (31.13% of affiliates, but only 17.44% of total employment given smaller average affiliate size of 385.79 employees) followed by Canada (13.03% of affiliates, 11.94% of total employment).

Table 2: Top 3 groups of counties of UBO in terms of percent of affiliates of foreign-owned enterprises and percent of employment

Country Group	Ranking	Percent of affiliates	Percent of employment	Average number of employees (Standard deviation)
Europe	1	44.28%	63.48%	987.21 (6,126.76)
Asia and Pacific	2	31.13%	17.44%	385.79 (1,930.61)
Canada	3	13.03%	11.94%	631.07 (4,734.33)

Industry of Affiliate

Table 3 shows the top 4 industry groups both in terms of the number of affiliates and in terms of total affiliate employment. In terms of the number of affiliates, the top 4 industry groups represent 74.01% of all affiliates. In terms of total employment, the top 4 industry groups represent 64.61% of all affiliate employment. Three sectors are included in both lists: Manufacturing (1st in terms of the number of affiliates and in terms of total employment); Wholesale trade (2nd in terms of the number of affiliates, 3rd in terms of total employment; and Financial activities (3rd in terms of number of affiliates, 4th in terms of total employment). Professional, scientific, and technical services is 4th in terms of the number of

affiliates but is not in the top 4 in terms of total affiliate employment, where it is replaced by Retail trade in 2nd.

Table 3: Top 4 industry groups in terms of percent of affiliates of foreign-owned enterprises and percent of employment

| Ranking | Ranking based upon: | | | |
| | Percent of total number of affiliates | | Percent of employment | |
	Industry	Percent	Industry	Percent
1	Manufacturing	28.82%	Manufacturing	36.86%
2	Wholesale trade	21.72%	Retail trade	10.74%
3	Financial activities	13.27%	Wholesale trade	9.17%
4	Professional, scientific, and technical services	10.20%	Financial activities	7.83%

Geographic Distribution of Affiliate Employment

Finally, while BEA inward FDI data are reported at the enterprise level, affiliates are required to break out their employment by state in which their employees are employed. Table 4 shows the top 10 states in terms of the percent of total affiliate employment as collected by BEA. The top 10 states contain 55.23% of all affiliate employment by state. Table 4 also compares the percent of total affiliate employment to the percent of all QCEW establishment employment within the same 10 states. Compared to all QCEW establishment employment, affiliate employment is more highly concentrated in New York and New Jersey, and less concentrated in California and Florida.

Table 4: Top 10 states in terms of percent of total employment

| State | US affiliate employment | | QCEW establishment employment | | Percent difference |
	Ranking	Percent	Ranking	Percent	
California	1	10.42%	1	11.28%	-0.85%
Texas	2	8.13%	2	8.08%	0.05%
New York	3	7.73%	3	6.46%	1.27%
Pennsylvania	4	4.84%	6	4.21%	0.63%
Illinois	5	4.54%	5	4.25%	0.29%
Florida	6	4.14%	4	5.60%	-1.46%
New Jersey	7	4.08%	11	2.85%	1.23%
Ohio	8	4.03%	7	3.79%	0.24%
Michigan	9	3.79%	8	2.96%	0.84%
North Carolina	10	3.52%	9	2.94%	0.58%

US Affiliates with Zero Employment

Finally, Table 5 shows the top 3 industry groups (4-digit NAICS) for the affiliates that report zero employment. These 3 industry groups account of 61.40% of affiliates that report zero employment.

These groups (Real estate, non-banking Holding companies, and Other financial investment activities) are all consistent with investments.

Table 5: The top 3 NAICS industry groups for US affiliates of foreign-owned enterprises with non-employment

Industry group (4-digit NAICS)	Percent of total affiliates
Real estate	54.01%
Holding companies, except bank holding companies	3.97%
Other financial investment activities and exchanges	3.42%

Section IV: Matching process

As noted in Handwerker and Mason (2013), EINs define business enterprises for tax purposes and are useful for linking establishments with their parents firms. While Elvery, Foster, Krizan, and Talan (2006) noted that most employers have just one EIN, EINs do not uniquely identify all firm; some firms may be comprised of multiple EINs. Handwerker and Mason (2013), for instance, show some evidence that some firms may use different EINs in different states. Additionally, firms involved in mergers and acquisitions have at least two EINs—at least one associated with the acquired firm and at least one associated with the acquiring firm. For those firms that have multiple EINs, a more fundamental problem is that those employees responsible for filing UI (the source of the QCEW) needn't be the same employees responsible for filing for other purposes, such as BEA mandated reports. For example, Handwerker and Mason (2013) show that EINs listed in SEC filings do not always appear in the QCEW.

Every establishment in the QCEW is associated with one EIN and each US affiliate for which establishments are to be linked has at least one EIN associated with it and often a listing of EINs are provided (for any enterprises that are consolidated within the report in which there is a 50% of greater ownership). Thus, the fundamental challenge in finding all of the establishments in the QCEW that comprise the US affiliates of foreign owned multinational enterprises is that of finding all of the establishment EINs for each affiliate. It is then a trivial step to combine the two data sources. This is done in a two-step process, first by automatically matching the EINs reported by each affiliates to all of the EINs in the QCEW, and second by manually reviewing these matches to remove mistakenly matched EINs and adding additional EINs that were not automatically matched.

Automated matching

The first step is automated. All of the reported BEA EINs are matched to those in the QCEW, and BEA reported affiliate employment is compared to that derived by summing employment for all of the QCEW establishments with matching EINs. This step is insufficient, as seen by assessing a simple metric that groups together matches that are "close" in terms of BEA and BLS employment; specifically, if derived BLS employment is within 20% of reported BEA affiliate employment. Only 39.59% of affiliates are considered to have "close" matches to the QCEW after automated matching, although these affiliates with "close" matches have 55.78% of all affiliate employment.

Analyst matching

After automated matching, a second time-intensive manual step is undertaken. In this step, analysts are assigned affiliates in order to review the automated matches. Assignments are made based on the quality of the automated match and the size of the affiliate: poorly matched, large affiliates are assigned first. An additional consideration is whether an affiliate is overmatched (BLS employment is greater than BEA) or under-matched (BEA employment is greater than BLS). This approach seeks to maximize limited analyst resources. Reviewing the automated matches is time-consuming. Further, the time varies depending on whether the automated match is overmatched or undermatched, with the former being less time-consuming. It is the case that the time spent reviewing an automated match increases by the size of the affiliate: on average there is a 0.68 minute increase (with a standard error of 0.06) per 1,000 employees. However, there is also considerable fixed time to review the automated matches, with 11.79 minutes taken on average for overmatched affiliates (with a standard error of 1.07 minutes) and with 20.86 minutes taken on average for undermatched affiliates (with a standard error of 0.54 minutes).

To facilitate the review, for each assignment the analyst is provided with measures of the quality of the automated match as well as information on the assigned affiliate and its matched establishments. The quality of the automated match is determined by the difference in affiliate employment and the sum of employment for the matched QCEW establishments. Affiliate information includes the affiliate name as well as any provided names of its consolidated subsidiaries, the address of the affiliate, the industry classification of the affiliate, a list of states in which it has employees, and a list of the matching EINs and matching establishments. For each matched establishment, in addition to the EIN associated with the establishment, the establishment's industry classification, establishment names and address data, and employment are also given.

The approach the analyst takes in reviewing the match depends in large part on whether the automated match is overmatched or undermatched. Overmatched affiliates are less time-consuming in part because all of the information necessary to determine invalidly matched establishments for the affiliate are already present. This is accomplished by comparing names, addresses, and industry classifications for the affiliate to those of the matched establishments. It is sometimes necessary to independently review outside resources via a web-search, the affiliate's website, any SEC 10-k filings made by the affiliate, or other resources such as Wikipedia or Bloomberg among many others. For undermatched affiliates, it is necessary to search for additional EINs via name or address matching (where additional names and addresses are determined via web-search or through reviewing SEC 10-k filings).

By removing mismatched EINs and adding any EINs that were not matched, it is generally the case that differences in BEA-provided affiliate employment and the sum of QCEW establishment can be reconciled. If extensive review of the affiliate shows the matched establishments to be correct and complete and BEA affiliate employment and derived QCEW establishment employment are close, the analyst signs off on the affiliate as being complete. It is sometimes the case, however, that despite extensive analyst review, it is not possible to reconcile BEA affiliate employment and derived employment for the matched QCEW establishments. In this case, the affiliate is recorded and sent to BEA for review. Another potential discrepancy is when the affiliate contracts with a Professional

Employment Organization (PEO) for all or part of its workforce. If the affiliate includes these employees as part of their total employment reported to BEA, it is not usually possible to match these employees in the QCEW establishment data. PEOs generally group together employees from multiple employer's establishments under one EIN.

Section V: Quality of Matches

Table 6 shows the number of EINs for each affiliate at three points in the matching process. Mentioned previously were the number of EINs that BEA provided for the affiliates that are being linked to establishments: 3.20 EINs on average (with a standard deviation of 14.69). Of these affiliates and their EINs provided, only 64.71% of the affiliates had a matching EIN in the QCEW after the automated matching procedures where run. On average, 1.90 (with a standard deviation of 3.69) EINs were found in the QCEW. Finally, to date 40.76% of the affiliates have been assigned to analysts to be reviewed. For the most part, the BEA-provided EINs that are in the QCEW are correct: only 0.06 (with a standard deviation of 0.58) EINs per affiliate are removed by the analyst. However, the EINs provided by BEA and in the QCEW are not complete. 1.28 (with a standard deviation of 5.35) EINs on average are added by analysts. After removing mismatched EINs and adding EINs that were not accounted for by the automated matching procedures, there are 3.62 (with a standard deviation of 7.53) EINs on average.

Table 6: Mean number of BEA provided EINs

Point in matching process	Percent of affiliates	Mean number of EINs (Standard deviation)	Mean number of removed EINS (Standard deviation)	Mean number of added EINS (Standard deviation)
Before automated matching: BEA provided data	100.00%	3.20 (14.69)	NA	NA
After automated matching: BLS matched data	64.71%	1.90 (3.69)	NA	NA
After analyst review: BLS matched data	40.76%	3.62 (7.53)	0.06 (0.58)	1.28 (5.35)

Tables 7 and 8 profile match quality at two different times in the matching process. Both tables include all of the affiliates that were matched to the QCEW. They both show the percent of the number of affiliates and percent of the sum of employment based on the size of the affiliate and the size of the difference between BEA reported employment and the sum of matching QCEW establishment employment (the formula is (BLS employment – BEA employment) / BEA employment). For affiliates for which no EINs were found, QCEW establishment employment is set to zero. (This is the case for the 35.29% of affiliates that were not automatically matched, and occurs occasionally after analyst review if the analysts are unable to find any establishments with valid employment for the affiliate.) Table 7 shows match quality immediately following the automated matching procedures.

Table 7: Match quality after the automated matching procedures

Over-matched: Sum of BLS establishment employment > reported BEA affiliate employment						
BEA Employment		Percent difference of BLS and BEA employment				
		<=10%	11-20%	21-50%	>50%	All
<=100	Number of Affiliates	39.15%	5.40%	7.55%	18.11%	70.21%
	BEA employment	1.98%	0.38%	0.46%	0.63%	3.46%
101-500	Number of affiliates	11.83%	1.93%	2.66%	1.39%	17.81%
	BEA employment	6.53%	1.07%	1.59%	0.69%	9.88%
501-1000	Number of affiliates	3.26%	0.72%	0.33%	0.48%	4.80%
	BEA employment	5.24%	1.18%	0.60%	0.82%	7.84%
>1000	Number of affiliates	4.71%	1.09%	0.91%	0.48%	7.18%
	BEA employment	55.36%	13.85%	6.93%	2.68%	78.83%
All	Number of affiliates	58.95%	9.15%	11.44%	20.46%	100.00%
	BEA employment	69.11%	16.49%	9.58%	4.82%	100.00%
Under-matched: Sum of BLS establishment employment < reported BEA affiliate employment						
BEA Employment		Percent difference of BLS and BEA employment				
		<=10%	11-20%	21-50%	>50%	All
<=100	Number of affiliates	5.22%	3.96%	6.54%	52.06%	67.78%
	BEA employment	0.29%	0.14%	0.18%	0.98%	1.59%
101-500	Number of affiliates	4.71%	1.61%	1.90%	7.97%	16.20%
	BEA employment	1.40%	0.48%	0.53%	2.18%	4.60%
501-1000	Number of affiliates	2.41%	0.76%	0.82%	2.24%	6.24%
	BEA employment	2.09%	0.63%	0.70%	1.83%	5.26%
>1000	Number of affiliates	3.25%	1.61%	1.72%	3.21%	9.79%
	BEA employment	30.06%	12.05%	15.42%	31.03%	88.56%
All	Number of affiliates	15.58%	7.95%	10.98%	65.48%	100.00%
	BEA employment	33.85%	13.30%	16.83%	36.02%	100.00%

Table 8 shows the improving match quality given the 40.76% of affiliates that were reviewed by analysts to date.

Table 8: Match quality given some analyst review

Over-matched: Sum establishment employment > reported BEA affiliate employment						
BEA Employment		Percent difference of BLS and BEA employment				
		<=10%	11-20%	21-50%	>50%	All
<=100	Number of affiliates	34.54%	4.97%	7.01%	19.71%	66.23%
	BEA employment	1.38%	0.29%	0.35%	0.61%	2.63%
101-500	Number of affiliates	11.75%	2.56%	3.00%	2.61%	19.92%
	BEA employment	5.00%	1.12%	1.32%	1.02%	8.47%
501-1000	Number of affiliates	3.35%	0.84%	0.56%	0.67%	5.43%
	BEA employment	4.11%	1.07%	0.79%	0.84%	6.81%
>1000	Number of affiliates	5.30%	1.23%	1.28%	0.61%	8.42%
	BEA employment	57.36%	11.98%	10.37%	2.39%	82.09%
All	Number of affiliates	54.94%	9.60%	11.85%	23.60%	100.00%
	BEA employment	67.85%	14.46%	12.83%	4.85%	100.00%
Under-matched: Sum establishment employment < reported BEA affiliate employment						
BEA Employment		Percent difference of BLS and BEA employment				
		<=10%	11-20%	21-50%	>50%	All
<=100	Number of affiliates	6.35%	4.73%	7.64%	51.73%	70.44%
	BEA employment	0.38%	0.19%	0.22%	0.88%	1.68%
101-500	Number of affiliates	6.43%	2.29%	1.93%	3.82%	14.46%
	BEA employment	2.04%	0.77%	0.54%	1.03%	4.37%
501-1000	Number of affiliates	3.38%	1.04%	0.64%	0.87%	5.93%
	BEA employment	3.07%	0.89%	0.61%	0.76%	5.33%
>1000	Number of affiliates	4.91%	2.38%	1.25%	0.62%	9.17%
	BEA employment	48.07%	23.81%	7.86%	8.89%	88.62%
All	Number of affiliates	21.07%	10.43%	11.45%	57.04%	100.00%
	BEA employment	53.56%	25.66%	9.23%	11.56%	100.00%

Table 9 provides a summary of Tables 7 and 8, classifying any match where BEA affiliate employment is within 20% of the sum of QCEW establishment employment. As is clear from the table, automated matching alone is insufficient, with only 39.6% of affiliates and 55.8% of employment being "close" matches. After reviewing 40.76% of the affiliates, most of which were large, the number of "closely" matched affiliates increased by only 5.91%, but total employment that is considered a "close" match nearly 24.50% of the time.

Table 9: Comparing BEA and BLS employment

Time of match	Percent	BLS employment is within 20% of BEA employment	BLS employment is **not** within 20% of BEA employment
After automated matching	Affiliates	39.59%	60.41%
	Employment	55.78%	44.22%
With analyst review	Affiliates	45.54%	54.46%
	Employment	80.28%	19.72%

Section VII: Opportunities for future work

We continue to research this topic, including plans to improve our matching methods and process. Additionally, we are planning publication possibilities.

We are looking into matching a subset of the 2013 BEA FDI data in order to develop ongoing cost estimates for the matching process. Determining the establishments that comprise the US affiliates of foreign-owned enterprises in 2012 is a very time-intensive process. However, matching establishments in ongoing years can leverage this work. Presumably the majority of affiliates will have very similar establishment composition one year to the next and the majority of analyst matching will only need to focus on new affiliates and affiliates that have seen merger and acquisitions.

Finally, we are examining the efforts necessary to link FDI-QCEW matched data with the occupational data in the Occupational Employment Statistics survey, similar to the work done by Handwerker, Kim, and Mason (2011). Using this link to bring in the OES data will allow us to examine the occupational profile of employment in foreign owned companies, and we think that will be a particularly valuable addition to the amount of detail already known about FDI in the US.

Conclusion

The BLS continues to leverage its investment in its QCEW Business Register by expanding its array of timely, accurate and relevant products.

References

Brinkley, Eden, Alistair Hamilton, and Therese Lalor, "Operationalizing Metadata Frameworks – An ABS perspective," Proceedings of the 4th International Conference of Establishment Surveys: International Conference of Establishment Surveys Montreal, Canada, June 11-24, 2012.

"International Economic Accounts," Bureau of Economic Analysis: https://www.bea.gov/about/pdf/international_fdius.pdf.

Elvery, Joel, Lucia Foster, C. J. Krizan, and David Talan, "Preliminary micro data results from the Business List Comparison Project," *Proceedings of the 2006 American Statistical Association Annual Meeting* (Alexandria, VA, American Statistical Association, 2006).

Farmer, Tracy E. and Michael A. Searson (1995), "Use of Administrative Records in the Bureau of Labor Statistics' Covered Employment and Wages (ES-202) Program," 1995 Bureau of the Census Annual Research Conference, Washington, DC, March 1995.

Handwerker, Elizabeth Weber, and Lowell G. Mason, "Linking firms with establishments in BLS microdata," Monthly Labor Review, June 2013, Vol. 135, No. 6.

Handwerker, Elizabeth Weber, and Lowell G. Mason, "Assessing the Usefulness of Census Bureau Multi-Establishment Data to Facilitate Linking Firms with Establishments in BLS Microdata," *Proceedings of the 2014 Joint Statistical Meetings.*

Handwerker, Elizabeth Weber, Mina M. Kim, and Lowell G. Mason, "Domestic employment in U.S.-based multinational companies," Monthly Labor Review, October 2011, Vol. 134, No. 10.

Robertson, Kenneth, Larry Huff, Gordon Mikkelson, Timothy Pivetz, and Alice Winkler, "Improvements in Record Linkage Processes for the Bureau of Labor Statistics' Business Establishment List." Proceedings for the 1997 Record Linkage Workshop and Exposition, pp. 212-221.

Searson, Michael A. and John Pinkos (1990), "The Bureau of Labor Statistics' Business Establishment List Improvement Project," 1990 Bureau of the Census Annual Research Conference, Washington, DC, March 1990.

Searson, M., "Strategies to Implement Electronic Collection of Multiple Worksite Report Data," Federal Committee on Statistical Policy, Washington, DC (November 2001).