Final Report for Small Business Administration Office of Advocacy Contract No. SBA-2113-AER-87

Point of view or opinions stated in this report are those of the authors and do not necessarily represent the position or policies of the Small Business Administration.

By

Joel Popkin and Company
1101 Vermont Avenue, N.W., Suite 201
Washington, D.C. 20005

February 22, 1991
# TABLE OF CONTENTS

**Executive Summary**

**Summary Report** .......................... 1

- Focus of Survey .................................. 4
- Survey Results - Employment .................... 7
- Survey Results - Proxy Branches ................ 18
- Government Affiliates ............................ 22
- Implications for Benchmarks ..................... 23
- Merger and Acquisition Activity ................ 26
- Observations and Recommendations .............. 27

**Appendix: Detailed Survey Information and Analysis** .......................... 33

- Introduction .................................... 33
- Preliminary Comparisons with Public Data Sources .................................. 34
- The Test Surveys .................................. 40
- The Full Validation Survey ....................... 45
- Employment Results from the Survey .............. 51
- Employment Deviations by Industry and Branch Type .................................. 74
- Double Counting of Employment ................... 76
- Subsidiary and Establishment Counts .............. 79
- Proxy Attributes .................................. 82
- Firms with Governmental Affiliations .............. 89
- Implications for Benchmarking USEEM to Other Data Sources ..................... 91
- Merger and Acquisition Activity ................ 96
Executive Summary

To support research required by the mission of the Small Business Administration's (SBA) Office of Advocacy, data for complete firms or enterprises are needed; data on an enterprise's individual establishments, the unit of observation for most government statistics, will not suffice. To meet the need for enterprise or firm data, the Office of Advocacy developed the U.S. Establishment and Enterprise Microdata (USEEM) file. This file, which is part of the SBA's Small Business Data Base (SBDB), is the main source of timely data on the employment size structure of U.S. firms and on other facets of small and large business. Frequently, USEEM is used to study business structure in specific industries (down to the 4-digit SIC level) and in detailed geographical areas (down to the county level).

The USEEM data base is constructed from a private data source, the Duns Market Identifier (DMI) file of the Dun and Bradstreet (D&B) Corporation. The need to use a private data base reflects the main stumbling block to the availability of Federal data on the size structure of business. It is the reality that most Federal collection programs now focus on obtaining data for the smallest possible units of companies -- establishments.¹ The SBA recasts the DMI data primarily to link establishments with their parent companies. SBA constantly reviews the resulting data and the methods used to develop them, with the objective of improving the sampling frame and the accuracy of the information obtained from it. It also is exploring alternative sources of enterprise data, including the possibility of developing a sampling frame from which the Federal government can directly collect data on small and large businesses and make such data accessible to users, as is done in Canada and other nations.

The research project reported here, and described in both a summary report and a detailed appendix, is one of those sponsored by SBA to validate and improve the USEEM file. The particular purpose of this research was to validate the data on the large business portion of the file. One reason for that focus reflects the seeming tendency for the number, and to a lesser extent the employment of large businesses to be overcounted in USEEM. A scientific sample was drawn of 480 firms with more than 500 employees in four major industries, almost 5 percent of the firms in USEEM with those size and industry characteristics. The firms

¹The Census Bureau's Enterprise Statistics (ES) is, apart from the SBDB, the only publicly available Federal source of data in which the establishments are linked and the enterprise can be analyzed as a unit. However, ES data are produced only every five years (and with a substantial lag) and customized tabulations of ES microdata cannot, for all practical purposes, be obtained. In virtually all other Federal business data series, the establishment or quasi-establishment is the reporting and analytical unit.
were sent questionnaires, approved by the Office of Management and Budget, designed to permit the comparison of the information in the USEEM files, developed from the data collected by D&B, with that reported directly by the firm. Almost half of the firms provided data on their firm's ownership and employment, broken down between the domestic and foreign work force. In addition, many firms provided more detailed structural information that could be used to make comparisons with that on the organizational units of the firms reported in USEEM.

Based on unweighted survey results, almost 60 percent of the employment numbers in USEEM matched (within 10 percent) those reported on the survey form after adjustments were made for all explainable deviations. Of the remaining firms, almost 30 percent showed fewer employees on their survey form than were reported in USEEM and slightly more than 10 percent showed more employees on their survey form than were reported in USEEM.

Analysis of these and more detailed survey results permit the adoption of a number of conclusions about the data on large firms in USEEM. First, when a parent reports more employment to D&B than can be obtained from summing that of its units appearing in the USEEM data base, the missing employment needed to reconcile the parent's report should continue to be imputed to proxy units (branches). In dividing the imputed employment among the proxy units, there is a tendency to overstate the number of employees per establishment and thereby understate the number of proxy establishments, although there are important industries where the opposite is true. After reviewing possible methods of improving estimates of the employment per establishment ratios used to determine the number of proxy units by using the other available information about the firm, clearly improved editing rules could be developed only in retail trade and public utilities.

Once created, the proxy establishments appear to be adequately assigned to an industrial classification, at least at the 2 digit SIC level. One exception is where an entire subsidiary is missing, in which case use of the distribution across industries of the rest of the company to allocate establishments sometimes leads to significant errors. The assignment of proxy establishments to geographical areas is frequently not accurate and may be a significant source of error in geographical analysis. However, this study validates as best practice SBA's decision to assign proxy establishments to geographical areas based on the geographical distribution of other branches in the data base.

One source of error in the assignment of employment to proxy units arises because the DMI file does not reflect the foreign-domestic split of employment for companies. If Dun and Bradstreet (D&B), DMI's producer, cannot or does not wish to obtain such data, other sources should be tapped such as the employment data from company reports filed with the Department of Commerce in the quinquennial census of overseas direct investment by U.S. firms or the related annual survey of direct investment abroad. If that is
not feasible, other possible sources include the 10-K reports that companies file with the Securities and Exchange Commission (SEC).

The survey and the following study of merger and acquisition activity did indicate some reasons for the overcount of large companies and the overcount of employment. The overcount of firms is larger percentage-wise than of employees. This implies that more firms are double counted than employees. There are some circumstances that can lead to an overcount of firms, without a concomitant double-counting of employees. One is if merger and acquisition activity exceeds divestitures and the DMI lags in capturing both, more, seemingly independent, firms will always be recorded in the file than actually exist. The determination of the relative frequency of total mergers and divestitures was not part of this study. However, a brief examination of acquisition and divestiture activity between 1986 and 1988 indicates that, while D&B does a relatively good job of tracking this type of information, some of the changes are not recorded in a timely fashion; these lags contribute to the overcount of firms. While it could not be proved from this sample, an additional source of this discrepancy could be that jointly held subsidiaries, especially those in which no owner holds as much as a 50 percent interest, may be counted as separate entities to a greater extent in DMI than by the Census Bureau.

Based on the survey analysis, the main factors that can contribute in USEEM to an overcount of employment but not of firms result from an interaction of D&B data collection methods with SBA editing procedures. When the sum of reported subsidiary employment is greater than reported firm employment, the editing rules increase firm employment to match the sum of subsidiary employment plus employment in the firm’s headquarters. Unfortunately, because employment at locations where both a headquarters and one or more subsidiary headquarters are located sometimes is double-counted, this procedure can result in more employment being assigned to a firm than should be. The survey data show that when the USEEM firm employment total has been changed, there is a definite improvement in the employment total only in about 13 percent of the cases. (A definite improvement is one where the DMI number is more than 10 percent different from survey employment while the USEEM number is less than 10 percent different from the survey.) About 28 percent of the time, making the change significantly worsens the employment estimates. The rest of the time it does not make a significant improvement in overall accuracy. The most straightforward way for the employment double counting problems to be resolved is for D&B to be aware of what circumstances result in such double counting and request a breakdown of employment by establishment when several of a company’s units reside at the same physical location.

The main circumstances that can lead to an overcount of both firms and employees is when a subsidiary of a company is not linked to it in the file, yet the parent reports the subsidiary's employment. This results in the employment being counted twice and two firms showing up in place of one. A case in which a holding
company did not fully consolidate the companies under its control would be one example of this problem.

The overall results of this research suggest that while some problems remain, SBA's persistent efforts to improve the methods it uses to transform DMI establishment records into USEEM enterprise data have resulted in a set of procedures that do a credible job of approximating the structure of large businesses. That is important because of the seemingly growing use of these data, as reflected in press reports, of studies about business structure and the reliance on these data in the preparation and publication of the President's annual report to Congress on the state of small business.

But SBA's efforts to improve USEEM are limited by the quality and timeliness of the data available to it in DMI. This executive summary contains some recommendations, discussed and justified in the detailed project report, for improvements that D&B could make in the DMI that would improve its suitability for the further development of USEEM. These include the collection of information on foreign employment and the improvement of the linkages it provides between companies and their subsidiaries and other establishments. And there are others. For example, the process of reconciling USEEM and government data could be facilitated if D&B solicited a more precise dating of the employment counts it obtains and used a more consistent definition of employment.

While such improvements are desirable from SBA's point of view, they may conflict with D&B's major purposes in collecting information about firms. Thus, the SBA should continue to explore alternative sources of enterprise data.
SUMMARY REPORT

The Small Business Administration's United States Establishment and Enterprise Microdata (USEEM) files have been constructed for each of the even numbered years from 1976 through 1988. These files are constructed from establishment data collected by Dun and Bradstreet for its Dun's Market Identifier (DMI) files. In producing USEEM, SBA assembles all of the establishments in the DMI that belong to each of the firms in the file. These firms can range from single establishment companies having only a few employees to giant multinationals with dozens of subsidiaries and thousands of employees.

The goal in constructing USEEM from Dun and Bradstreet's data is to produce a timely representation of the size and industry structure of the U.S. private business community and an accurate measure of its employment by firm size. SBA constructs its own data base because comprehensive business data identified by firm size, to the extent it exists within the federal government, is not made available to SBA, particularly below the national level. However, even access to the government's most complete business data base would not supply SBA with comparable information on all the aspects of each firm that are measured in USEEM. USEEM is frequently used to study business structure in specific industries (down to the 4-digit SIC level) and in detailed geographical areas.
The purpose of this project has been to evaluate certain aspects of the D&B data base SBA uses to construct USEEM and of the methods SBA uses to remedy certain deficiencies in those data. These include missing establishments, the inclusion of foreign employment and employment in public agencies.

To do that evaluation, several questions had to be answered about both the basic DMI data and the methods SBA uses to edit and to refine those data. Some of those questions are presented below.

1. Are the individual establishment employment counts, industry classifications and locations accurate in DMI?

2. Are the firm employment counts in DMI accurate for domestic employment, for domestic plus foreign employment or not accurate for either?

3. If the DMI firm employment is not accurate, can anything be done to make it more accurate? Specifically, have the methods SBA has used to create a better estimate of "true" firm employment in USEEM, produced a more accurate estimate and if not, are the USEEM estimates biased either up or down?
4. Are all the establishments of every firm in the DMI database? If not, should SBA continue its current practice of representing those missing pieces by imputing proxy branches? If all the pieces of each firm are indeed in the database, but the information necessary to link each establishment to its firm is sometimes missing, does improving the representation of each firm actually make the aggregated data base of all the firms less accurate?

5. If the establishments missing from the firm's structure are going to be represented by proxy branches, how should the employment be assigned to them and how should SBA split that employment between foreign and domestic establishments? Is SBA's current method of ensuring a match at the broad industry level between the total amount of employment assigned to foreign subsidiaries in USEEM and BEA's foreign employment counts sufficient or must it be correct for each individual firm?

6. What set of "employment-per-establishment" numbers will result in approximately the correct number of proxy establishments being created, so that not only is the employment total correct, but also the organizational structure is close to right? SBA currently uses the size, industry and location of the firm to estimate the employees-per-establishment number used to determine the
number of proxy establishments assigned to each firm. Does that methodology result in approximately the correct number of establishments being created?

7. How should proxy establishments be classified by industry and assigned a geographical location? Is SBA’s current method of using the industry classification and location of the headquarters that is missing establishments an accurate representation of the missing pieces?

8. Should every firm in the DMI data base be included in the USEEM data base? Is there any way for SBA to separate the enterprises with governmental affiliations from the private firms in the same industrial classification?

9. Are mergers, acquisitions and divestitures promptly and properly recorded in the DMI and USEEM?

Selecting the Focus of the Study

The process of validating the entire data base and answering these questions for every establishment in it was much too large an undertaking for a project of this size. Instead this summary report and the detailed appendix cover the information collected about a subset of firms in the data base and the questions that could be answered about them. Since many of the questions address
issues in the linking of establishment records to create entire multi-establishment firms and the creation of proxy branches to represent missing establishment in the firm, firms that had substantial employment in proxy branches were a reasonable subset to study. The number of proxy branches in the 1986 USEEM\(^2\) file and the amount of employment in those proxy branches were examined to determine which sectors of the data base would be the most appropriate ones to cover in the study.

Looking at the distribution of proxies by industry and firm size, it was discovered that only about one-third of all the proxy branches were in businesses with more than 500 employees. However, those branches accounted for more than 70 percent of the employment assigned to proxy branches.\(^3\) In addition, of the five industries that accounted for almost 96 percent of large business employment, four had more than twenty percent of their employment in proxies. Those industries were 1) manufacturing, 2) services, 3) retail trade, and 4) transportation, communications and public utilities (TCPU). Consequently, this study focused on the multi-establishment firms with 500 or more employees in the manufacturing, retail trade, services, and transportation, communications and public utility industries.

\(^2\)The 1988 USEEM file has only been completed within the past few months and was not available for the survey comparisons.

\(^3\)Large firms accounted for about 30 percent of domestic proxy branches and 63 percent of domestic employment in proxy branches.
During the early part of this project several different procedures were explored to see which method was best to answer the questions posed above. Those included using public data sources and testing two surveys. After comparing the information gathered from each source, it was decided that a survey was the best method available to provide the widest scope of information needed to answer the questions.

The sample for the survey consisted of 480 large firms from the 1986 USEEM file evenly divided among the four industries included in the study. These 480 firms were about 5 percent of the firms in USEEM with the desired size and industry attributes. Almost half of the firms (46 percent) provided a response that was usable in some form.

The survey provided a lot of information; however, it could not explicitly answer all of the questions posed. The first question about the accuracy of the data reported about each DMI establishment was one that could not be answered in the scope of this study. While some establishment detail was collected, most of the information from the survey pertains to more aggregate levels of the firm. The establishment employment totals that were reported in the survey were sometimes quite close to those in DMI, but in cases where there were deviations, it was not always clear whether the DMI record was in error or whether consolidation of more than one establishment was taking place either in the DMI or
in the information collected in the survey. Consequently, a definitive answer to questions pertaining to the overall accuracy of the individual establishment records was not possible from the data collected from this survey.

Survey Results - Firm Employment

The determination of the accuracy of the firm employment counts in the original Dun and Bradstreet DMI records was more definitive. A relatively large percentage (58%) of the firm employment counts reported on the DMI record are accurate to within ±10 percent of the domestic firm employment reported on our survey. Some of the firms have foreign employees included in the DMI count; if that employment was removed from the DMI total, the percentage within a 10 percent range of error increased slightly. However, as will be discussed later, it is not always clear how many foreign employees should be removed.

In the majority of cases the firm employment total on the DMI is the firm employment number that is carried over into the USEEM files. Frequently, however, the editing procedures used to construct USEEM may result in the USEEM firm employment total being

---

"For each firm the percentage difference between total firm employment as reported on the DMI record and domestic employment as reported on the survey was calculated. Each firm was then assigned to a group based on that percentage deviation between the two employment numbers. The percent of all firms that fall into each group is shown on the following charts."
different from the DMI firm employment total. There are cases, of course, where the DMI total is not correct and there are cases where the DMI record does not report any firm employment number. When firm employment does not exist in the DMI, using the current method of estimating a firm employment number probably improves the accuracy of the data base, at least for the large firms being studied in this analysis.

It is less clear that in other instances the current method of adjusting the DMI firm employment total to a new USEEM firm employment number always improves the employment estimate. The distribution of the deviations between DMI firm employment and the domestic firm employment total from the survey and between 1986 USEEM firm employment and that of the survey are shown in (Chart 1). A smaller percentage of the firms (55%) fell within the ±10 percent boundaries when the USEEM number was used than when the DMI number was used (the 58% mentioned above). This implies the DMI firm employment total is more accurate than the USEEM firm employment numbers.

5The 55 percent is based on the 194 surveys useable for this analysis. Thirty-two percent of the firms in this group were from the TCPU industry (which makes up only 7.5 percent of large firms in the four industries in USEEM being used in this analysis). Since TCPU also showed a larger percentage of firms in the ±10 percent category than did the other three industries, a distribution weighted by relative industry size would show about 50 percent of the firms in the ±10 percent category. The comparable weighted number for the DMI distribution would be approximately 53 percent.
Comparison of 1986 USEEM and DMI firm employment to survey employment

- USEEM, DMI less than survey
- USEEM, DMI greater than survey

Percent employment deviated from survey:
- USEEM
- DMI

Categories: <100, 76-100, 51-75, 26-50, 11-25, 6-10, +5, 6-10, 11-25, 26-50, 51-75, 76-100, >100
The most common reason for the USEEM firm employment number to differ from the DMI employment number is that when the employment in the different pieces of the firm is summed up, it is greater than the employment total for the firm shown on the DMI record. When that happens, the USEEM firm employment number is set equal to the sum of the employment in the pieces of the firm.

Based on the survey data only about 13 percent of the time when the USEEM differs from the DMI is there a definite improvement in the firm employment total. (A definite improvement is one where the DMI number is more than 10 percent different from survey employment while the USEEM number is less than 10 percent different from the survey.) About 28 percent of the time making the change significantly worsens the employment estimate (moving the USEEM number outside the ±10 percent boundaries when the DMI number was within the 10 percent boundaries). The rest of the time it does not make a significant improvement in overall accuracy, either because the USEEM and DMI deviations both fall within the 10 percent boundaries or because they both fall outside the 10 percent boundaries.

One reason that adjusting the employment number sometimes worsens the employment estimate is that the most common method of making the adjustment may inadvertently double count employees. The method of summing the subsidiary employment totals and adding that to the employment in the headquarters of the ultimate owner
and the employment in the branches that report directly to that headquarters frequently results in at least the headquarters employment being counted more than once. This happens because the headquarters of the ultimate owner and some of its subsidiaries are often at the same address. The DMI record for each of those establishments frequently shows the total employment at that address as the establishment employment for each one. Therefore, when the totals are summed the headquarters gets added in more than once. This can also happen when the headquarters of two subsidiaries are at the same address.

Several other types of double counting (detailed in Table 1) were uncovered in the analyses of the 1986 data. Some of them caused proxy branches to be created during the USEEM editing process and some did not. While the procedures for adjusting the DMI total to produce the USEEM total do include some control checks to prevent this problem, they should be reexamined in light of the specific double counting problems identified to see if further refinement could improve them. All of these problems could be minimized by relatively minor changes in D&B recording and collection practices. Some of the double counting could be caught by SBA through a method of checking the employment in establishments sharing the same address against each other. However, the logistics of that comparison could be very costly.
Table 1: Types of Employment Double Counting Observed in the Surveys

1. Two subsidiaries at the same address are each assigned the same total employment. There is only enough branch employment to fill one of them and proxy branches are created to fill the other. (This may be a problem of a subsidiary that reports to another subsidiary with the links between them missing.)

*D&B could improve subsidiary to subsidiary linkages and more carefully separate employment between the different establishments sharing an address. SBA could cross check the employment of different establishments at the same address.*

2. Establishments of a subsidiary aren't linked to it, but are shown as separate subsidiaries instead. Because all the establishments of that subsidiary can't be found proxy branches are created, but the employment is actually already within the firm structure.

*D&B could improve linkages.*

3. A) Two establishments with different names at the same address are assigned the same establishment employment. This doesn't create new proxies, but does double count the employment of the establishment. This happens most frequently when a subsidiary and ultimate owner headquarters are at the same address and each is assigned all the employees at that address.

*D&B could more carefully separate employment between different establishments sharing an address. SBA could cross check the employment of different establishments at the same address.*

B) One establishment is assigned two different Duns numbers by D&B and is counted as two separate establishments.

*D&B could more carefully determine if an establishment is already in the DMI. SBA could cross check establishments at the same address.*

C) One establishment is assigned two different Duns numbers one at a P.O. Box and one at a street address.

*D&B could determine both the street address and the P.O. Box of each establishment and check for both in the DMI records before assigning a new Duns number.*
4. The same subsidiary headquarters will be assigned two different Duns numbers, one at a street address and one at a P.O. Box. Since there are only enough branches to fill one of the subsidiaries, the other is filled with proxy branches. D&B could determine both the street address and the P.O. Box of the headquarters establishment and check for both in the DMI records before assigning a new Duns number.

5. One establishment is being moved over a period of time from one place to another. Both addresses are operating for the amount of time it takes to move and both get assigned the total employment for the establishment. D&B could control this; however, it is probably a relatively short term problem provided the establishment record is updated frequently enough to show the final closure of the establishment at the first address.

Proxy branches are created to fill in for pieces of the firm that can not be located in the DMI files. It can not be determined if these establishments can not be located because they are not on the DMI data base or because the linkages to their owners are not in place. It is very unlikely that all the missing establishments are actually on the DMI but, when comparing establishments on the data base to survey data some subsidiary pointers to other subsidiaries appear to be missing. However, short of searching the data base for all the missing tops, subsidiaries and establishments identified in the survey, it is not possible to determine if all the missing parts of a firm are truly missing or only misplaced. If missing establishments are in the data base but not linked to their owner they are probably not being included in the proper firm size class. If some of the missing tops or

6The parent Duns numbers were not included in the extract JPC used. Therefore this can only be deduced from total subsidiary employment and the number of proxies produced.
subsidiaries are there, they may be contributing to the overcount of large firms that plague USEEM. But, without knowing if the missing pieces are there, it is impossible to determine if creating proxies to fill in missing firm employment actually results in double counting it, thus making the entire data base less accurate as it makes the individual firm more accurate.

It does seem clear that for some of the firms studied there were pieces of the firm missing from the USEEM representation of the firm. When the employment assigned to the proxy branches associated with each firm in the sample was removed from the USEEM firm employment total and compared to the domestic firm employment from the survey, the USEEM number tended to fall significantly short of the survey totals. While about 55 percent of the firms in the sample were within ±10 percent of the survey totals before the employment in the proxy branches was removed, as can be seen in Chart 2, only 20 percent were within those boundaries after it was removed. Therefore, at least some of the employment presently being assigned to proxy branches in large firms does need to be there to create a more accurate representation of the firm. The employment in proxy branches can not be ignored.

However, determining how much employment should be assigned to proxy branches is not always straightforward. While subtracting the employment in the parts of the firm that are in the data base from the total firm employment seems to be the logical way of
2: COMPARISON OF 1986 USEEM AND USEEM LESS PROXY EMPLOYMENT TO SURVEY EMPLOYMENT

PERCENT OF FIRMS

USEEM LESS THAN SURVEY

USEEM GREATER THAN SURVEY

PERCENT EMPLOYMENT DEVIATED FROM SURVEY

□ USEEM □ USEEM LESS PROXIES
proceeding, it does not always result in the right number. As was mentioned earlier there are certain types of double counting that result in too many proxies being produced. There is another type of double counting that results in too few proxies being produced.\(^7\) In addition, it is not always clear whether or not the total employment number on DMI and USEEM includes a firm's foreign employees. Sometimes there are also problems with seasonal employment and the definition of firm employment.

If more care is taken in controlling the double counting problems, and D&B institutes a more consistent definition of firm employment, then subtracting what is there from the total will probably give a relatively good measure of how much employment needs to be assigned to proxy branches.

**Foreign Employment**

Once the total number of employees to be assigned to proxies is determined, those employees then need to be split between foreign and domestic proxy branches. *Ideally D&B would collect consistent domestic and foreign employment counts for each firm.* While they do identify a large number of multinational firms for

\(^7\)This happens when the total firm employment number is greater than the sum of the pieces that are present either because a subsidiary is missing or there is foreign employment. All of this "missing" employment should be assigned to proxy branches, but the double counting of the headquarters stands in for part of the "missing" employment and too few proxies are created.
SBA it is apparently not possible for them to determine the total number of foreign employees included in each firm total. Since the procedures for assembling USEEM guarantee the total foreign employment number for each broad industry group in USEEM matches those in the BEA survey of U.S. foreign investment, the total foreign employment in the data base is probably about right. However, if the individual firm records do not accurately portray their foreign employment, a decision must be made as to whether accuracy for broad industry groups is good enough for the uses to which the data are put or whether the inaccuracies at the firm level will severely impact micro analysis being done with the data.

There are some inaccuracies that occur from using the current methodology for handling foreign employment. While the methodology does seem to do a reasonable job of identifying firms that have a large number of foreign employees, it does not do as good a job of estimating how many foreign employees a given firm will have. But, it is almost equally likely to assign too few foreign employees as too many foreign employees to a firm. Twenty-two firms in the survey reported having foreign employees. Twelve of those had been assigned foreign proxies in USEEM and of the remaining ten, four had between 100-250 foreign employees and the remaining six had an average of only 24 foreign employees. The current methodology does

---

There is a question about whether the foreign employment definition used by BEA matches the one firms are reporting to D&B. That question can not be answered by this study.

15
seem to correctly pick out the employers with large numbers of foreign employees.

Of the seven firms that were assigned foreign employees but did not report any, four were assigned 900 or more employees.⁹ For those individual firms, that assignment of foreign proxies resulted in a significant error in their domestic structure. Even for firms with foreign employees, the foreign employment assigned was frequently so far from their actual employment that it made a substantial change to the structure of the firm. Only one company showed foreign employment that was within 10 percent of the number USEEM had assigned, and the remaining firms showed deviations averaging almost 200 percent. Chart 3 shows the effects of the individual firm inaccuracies on the error distribution of the survey respondents. With the employment in foreign proxies removed, only 53 percent of the firms fall within the ±10 percent boundaries compared to 55 percent for USEEM including all the proxies. This is despite the fact that 22 firms in the survey reported foreign employment that needed to be removed from the USEEM totals to make them definitionally match the domestic employment number from the survey.

⁹It is impossible to determine if these seven firms were from the D&B list of multinationals or if these firms were chosen randomly to bring the foreign employment levels up to the BEA benchmark totals.
3: COMPARISON OF 1986 USEEM AND USEEM LESS FOREIGN PROXIES TO SURVEY EMPLOYMENT

PERCENT EMPLOYMENT DEVIATED FROM SURVEY

- USEEM
- USEEM LESS FOREIGN PROXIES

PERCENT OF FIRMS

USEEM LESS THAN SURVEY

USEEM GREATER THAN SURVEY

<100 76-100 51-75 26-50 11-25 6-10 6-10 11-25 26-50 51-75 76-100 >100
Since the accuracy of some firm records is certainly reduced by the current method of assigning foreign employment, a cost-benefit decision will have to be made. There does not appear to be an easy way for SBA to fix this problem. The most straightforward way would be for D&B to separate domestic and foreign employment when it does the initial data collection. If that is not possible, information from BEA's survey of direct investment abroad could provide SBA with the firm specific employment totals (of course, only in their benchmark years do they collect data from all domestic firms with foreign affiliates). If the confidentiality problems with the BEA survey could not be solved, there are other methods that would improve the data, but they would be time consuming to implement. Using either the D&B list augmented by other sources such as information about foreign affiliates from Mergers and Acquisitions or BEA's list of multinational firms (if they would provide it) the SBA could either check the 10-K reports for those firms (which frequently provide a foreign employment total) or survey those firms and ask them about their foreign employment. If the SBA did not wish to do this every two years, perhaps one data collection could provide the information with which to construct relationships between foreign employment totals and other firm characteristics that could then be used to better predict how many foreign employees should be assigned to each firm.

10 The SBA might be able to use the methodology from this survey and ask the firm to provide a copy of the last survey it filled out for BEA.
Attributes of Proxy Branches

Determining the ideal size for each proxy establishment is not a simple process. It is one SBA has been trying to refine for some time. The proxy branches are being created not only to assure that the full firm employment is being included in establishment counts, but also to approximate the structure of the missing portions of the firm; therefore, having about the right number of proxies is almost as desirable as having the right amount of employment included in them. Unfortunately, the burden of providing full firm structure was more than most of the survey respondents wished to undertake. Therefore, the information about the appropriate size and number of the establishments being represented by the proxy branches is limited.

About 35 Census Report of Organization forms were received with the surveys, although not all of those could be used to do the proxy branch analysis. These forms provided the best measure of what the appropriate number and size of the proxy branches should be. In addition to the Census forms, the survey asked for a 1986 count of each firm’s retail branch locations. Those counts were used to determine if the correct number of proxy branches was being created for the retail firms whose USEEM and survey firm employment numbers were quite close. From the proxy branch counts it was
possible to infer whether the correct size of proxy branch was being used.

In general, the number and size of the proxy branches that could be checked were not correct. Based on the few Census forms available, the proxies tended to be too large for the industries overall although no attempt was made to weight the responses by the relative size of the industries. In retail trade, where the best information was available, the proxies were not consistently either too large or too small nor were there consistently too many or too few of them. Proxy branches tended to be larger than the missing branches for manufacturing and TCPU, whereas services and retail trade were split almost equally between those that were too large and those that were too small. The sample for each of these industries was very small and must be considered in that light.

Because the retail branch counts were available from the survey form, there were 20 more firms in retail trade that could be examined.\footnote{These firms were chosen from the set of firms that provided retail branch counts, but not Census forms. Only the firms in this group whose USEEM employment was within 10 percent of the survey employment were included in the analysis.} Of those, 1 had almost exactly the same number of proxy branches as missing branch establishments. Eleven firms had more proxy branches than were missing implying the average proxy branch was too small and 8 had fewer proxy branches than were needed implying the average size was too large. The retail firms
that did provide Census forms showed a very similar pattern with slightly more than half showing proxy branches that were too small.

There is not a clear way to improve on the current methodology for producing the employee per branch averages used to determine the size and number of the proxies. Our sample size was not large enough to analyze the use of establishments per firm as an alternative to employees per establishment in determining the number of proxies other than to comment that a wide variation in establishments per firm was observed among the sample being examined. An analysis was done to determine if the size of the proxies could be based on the average size of the non-headquarters establishments present in USEEM and associated with the same headquarters where the proxies were being created. However, as has been noted before in the USEEM literature, the establishments in the data base are not always representative of the ones that are missing. In less than half the cases would using the other establishments have been a substantially better indicator of the correct size of the proxies than the current method. Using the other establishments works better for the retail trade and public utilities industries, where the branches tend to be very uniform. The employment size of a McDonald’s outlet, for example, does not have an extremely large variance.

The industry classification to be assigned to each proxy branch is relatively straightforward when branch locations are
missing from well defined businesses. Assigning the industry classification of the headquarters the proxy reports to is logical and on the whole seems to be accurate at least at the two digit SIC level. Problems arise when entire subsidiaries are missing. Then the SIC codes of what is in the data base may not accurately reflect what is missing.

Assigning proxy branches to correct geographical areas is more difficult. "Roughly half of the firms looked at had proxies in the wrong state." Currently the geographical location of proxy branches is also currently assigned based on the location of the headquarters to which the proxy is assigned. However, this obviously does not work as well as the industry classification does. For firms that operate in a concentrated geographical area, using the location of the headquarters will generally put the proxy branches in the right state. However, for more dispersed firms the proxy branches often end up in the wrong state using this method. Using the location of the actual establishments that are in the data base would be a better method. While this will not always give the correct geographical distribution, it does tend to be better in many cases than the current method and is almost never worse. Unfortunately, the locations of the establishments missing from the firm do not seem to be random. Often most or all the establishments in a certain state are not there. Sometimes this is because a subsidiary is missing, but that is not obviously the case in other instances. However, distributing the proxies equally
among the states already represented in the data base does not consistently overcome this problem; therefore, the proxies should be distributed according to the number of actual branches in each state or the employment in actual branches in each state. Both methods tend to give similar answers.

Businesses Affiliated with Governmental Bodies

It has been previously recognized by SBA that enterprises associated with governmental bodies but conducting businesses similar to private businesses are very difficult to remove from USEEM, and therefore, increase the employment totals above the private industry employment benchmarks with which USEEM is being compared. The survey confirms that this is a problem, but can only recommend minor adjustments to the current procedures.

Currently all large employers in SIC 8211 are removed because they are assumed to be public school systems. Both of the employers in SIC 8299 that were on the survey list were also public schools. This SIC should be checked for other misclassified institutions. Large employers in SIC 8222 could be added to the list of SICs removed since these are generally community college systems. The large employers in SIC 8221 are somewhat harder to classify although about half of these in the survey sample were also state institutions. There are readily available listings of colleges and universities that identify whether these are public or
private institutions. Therefore, it would be possible to make up a one-time list (a relatively small number of colleges and universities would fall into the new births category each year) that identifies each of the employers in this group as being public or private and removing the public employment from the data base.

Beyond the education sector identifying public sector employees becomes much more difficult. There are obviously employers in TCPU that fall into this category and some of the survey respondents explicitly identified themselves as having governmental affiliations. However, they do not fall into just one SIC or make up the predominate part of particular SICs. Therefore, they can not be separated from private industry employers as easily as the educational institutions can.

Implications for USEEM Benchmarks

One of the main uses of USEEM and its USEEM subset is to provide estimates of job growth by firm employment size and industry. To provide accurate estimates requires that USEEM approximates the size and structure of the U.S. economy relatively well. To determine if USEEM is fulfilling that requirement it is generally compared or benchmarked to other collections of data that delineate the U.S. private business community.
The most accurate benchmark should be total employment in domestic establishments in the private U.S. economy. However, as the comparisons are made at a more detailed level by size class, industry and geographic region definitional differences between different data sources, the problems of identifying and classifying firms and assigning employment to industries, and the inclusion of foreign employment as well as the USEEM editing procedures themselves can begin to affect benchmark comparisons. This study cannot determine the accuracy of the overall data base, but can provide insights into why counts of large firms and their employees in USEEM do not always match those in other government sources.

While the survey was designed to examine the accuracy of the groups of records that have been combined to form large firms, it also provided some basis for determining why USEEM's large business employment might differ from the benchmarks it is being compared against. While Employment and Earnings and County Business Patterns are used to benchmark total establishment employment, the Census Bureau's quinquennial Enterprise Statistics is the only published source of government data available with which to compare firm employment by firm employment size class.

12Using the employment counts from BLS's Employment and Earnings and the Census Bureau's County Business Patterns as benchmarks, USEEM's private, nonagricultural, domestic employment counts are almost 9 percent above the BLS and CBP counts for 1986. That difference drops to about 6 percent once an estimate of government affiliated employment is removed.
Table 2 summarizes several reasons why the USEEM employment totals for large business may differ from the Census data (for most industries the 1987 Enterprise Statistics shows fewer large businesses and fewer employees in large businesses than the 1986

### Table 2: USEEM AND CENSUS BENCHMARKS

<table>
<thead>
<tr>
<th>Circumstances that might increase USEEM total employment above Census total for 500+ category</th>
<th>Circumstances that might make USEEM total employment less than Census for 500+ category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Double counting of headquarters or other establishments.</td>
<td>1. Missing tops provided they are not just improperly linked.</td>
</tr>
<tr>
<td>2. Not enough foreign employment removed.</td>
<td>2. Too much foreign employment removed.</td>
</tr>
<tr>
<td>(Each of these is seen in individual firms, but the effect on total employment in the 500+ category can’t be determined)</td>
<td></td>
</tr>
<tr>
<td>3. Inclusion of government enterprises.</td>
<td>3. Subsidiaries that aren’t linked to parent and are too small to be in 500+ category without that link. Of course, if the parent’s total employment number includes the subsidiary employment then it would be proxied.</td>
</tr>
<tr>
<td>4. DMI picks up a flow measure of employment rather than a stock measure of employment.</td>
<td></td>
</tr>
<tr>
<td>5. Subsidiaries that are large enough to fall into 500+ group, are not linked to parent, but parent reports their employment. This entity would then be proxied as well as exist separately.</td>
<td></td>
</tr>
</tbody>
</table>
USEEM files). Some of these have been directly observed from the analysis of the survey data and some have been inferred as likely possibilities.

**Merger and Acquisition Activity**

A brief study of merger and acquisition activity (p. 95) showed that there are times when a company is not properly linked to its new owner. Of the 31 instances of merger, acquisition or divestiture activity in the surveyed firms during the 1986-88 period, 24 were correctly reported in the 1988 DMI files and most of the ones that were not reported took place during 1988. In cases where the change was not correctly reported it was not possible to determine if a firm's employment was being reported both by it and its new owner. A further study of merger and acquisition activity involving partially owned firms identified three firms out of 66 that were probably contributing to an overcount of large firms in the 1988 USEEM file. There were also some instances in which small subsidiaries were not linked to their owning firm; that would contribute to an overcount of small firms. The identification of other cases of establishments being counted twice would require a search of the data base for all the pieces of a firm that are being proxied. However, incidences such as these could account for the employment and firm overcounts in USEEM.
In general, it would be expected that Enterprise Statistics would have fewer linkage problems than D&B data. The main focus of Census when it creates the Enterprise Statistics data is to reconcile entire firms at one time using consistent domestic employment counts as of a specific date. D&B's major focus is collecting credit data on the total firm; therefore, the collection of the data for the establishments of a firm is of lesser importance and takes place over an extended period. Consequently, the establishments are not reconciled to the firm level until USEEM editing is completed.

Observations and Recommendations

In short our observations and recommendations include the following items:

A. Current procedures that should be continued.

1. The employment currently assigned to proxy branches should continue to be included in the data base and used at both macro and micro levels.

2. Proxy branches do seem to be standing in for more than one missing branch location; therefore, the creation of multiple proxy branches is reasonable.
3. The current use of the industry classification of the headquarters as the SIC for all of the proxy branches works relatively well and should continue.

B. Changes to current USEEM procedures that are important in improving macro employment counts.

1. The controls used to prevent large increases in employment when adjusting the DMI to the USEEM level should be reexamined in light of the types of employment double counting identified in the survey. The current controls are not preventing all double counting (which probably can not be done) and while they may be controlling the number of unnecessary proxies being created they are not preventing it. Cross checking employment in the establishments at the same address is one way of combating this problem especially for firms where the USEEM and DMI firm employment numbers will be substantially different. (Not all branches at the same address can be considered double counts, two establishments at the same address in and of itself does not mean that double counting is taking place.)

C. Changes to current USEEM procedures that would improve the data base and probably would not require major expenditures.
1. The proxy branches should be distributed among all the establishment locations associated with a headquarters on the USEEM rather than being assigned only to the headquarters location.

2. SIC 8299 should be examined for other large schools and those found should be removed from the data base. Large employers in SIC 8222 should be removed from the data base for the same reason large schools are removed. Large employers in SIC 8221 should be identified as public or private institutions and the public employment removed. The large employers in SIC 8651 should be examined. At least one of those employers is a government agency that should be moved to another SIC.

D. Problems that would require more extensive resources or need further study.

1. Some of the proxy branches are clearly standing in for foreign employment; however, assuming most of the proxies of large multinationals are foreign proxies often results in assigning too many foreign employees to the firm. Better estimates of foreign employment by firm might be obtained from BEA, 10-K or survey information. The resource expenditure to collect these data is probably
worthwhile if the data are being used extensively for micro analysis.

2. Once the employment to be put in proxies is identified, the number and size of proxies being created is generally not correct although the branches are small enough that a difference of 5 or 10 employees caused a large percentage error. There is a tendency for the proxy branches to be too large but that varies by industry. Unfortunately, a clear way to improve the current methodology was not found. Using other data from the firm structure does not consistently provide a clear improvement in employee per establishment estimates for the proxy branches, although it might be useful for retail trade and public utility offices.

3. It is still not clear why there are more large firms and large firm employment in USEEM than in Enterprise Statistics. However, it is clear that not all mergers and acquisitions are promptly recorded in the data base. This leaves two firms where there should only be one and probably contributes to firm and employment overcounts. Timing differences between when the merger takes place and when all the partners in the transaction are interviewed by D&B mean that a certain number of firms will always fall into this category because not all the
firms will be interviewed before the file SBA uses is closed. There was one case where the linkages in the DMI tenuously pointed to the correct owner through a secondary pointer. But, the USEEM editing process did not maintain that connection. This indicates that in cases where pointers are inconsistent in the DMI, the current hierarchy for making them consistent in USEEM may result in some linkages being lost.

E. Changes Dun and Bradstreet should be encouraged to make in their data collection.

1. Make efforts not to assign a new Duns number to a branch that already has a Duns number.

2. Make efforts not to assign different Duns numbers to branches, subsidiaries or headquarters at both their street address and P.O. Box number.

3. Be more careful about separating the employment at headquarters locations among the different units of the firm at that address.

4. Improve subsidiary to subsidiary pointers.
5. Improve the identification of firms that are including foreign employment in their employment counts and perhaps collect separate foreign and domestic employment totals.

6. D&B should establish and use a standard, consistent employment definition such as number employed in the month of March or in the month the D&B survey is being filled out. It should establish a consistent treatment of part-time workers. A consistent definition would help ensure that the firm is providing a measure of employment at a given point in time rather than providing a count of employees over an extended period of time. While it can not be proven from the survey that D&B sometimes is given W-2 employment counts by firms, that would be one explanation for large overcounts especially in the retail food firms where turnover is quite high.
APPENDIX: Detailed Survey Information and Analysis

Introduction

Under the Small Business Policy Act of 1980, the U.S. Small Business Administration has been charged with "establishing and maintaining an external small business economic data base for the purpose of providing the Congress and the Administration information on the economic condition and the expansion and contraction of the small business sector ..." The result of that requirement is the United States Establishment and Enterprise Microdata (USEEM) file and a subset of those records called the U.S. Establishment Longitudinal Microdata (USELM) file.

For each even numbered year between 1976 and 1988, SBA has constructed the USEEM files from data collected by Dun and Bradstreet for its own Dun's Market Identifier (DMI) files. Most of the information for the DMI is collected from establishments because D&B has received a request for credit information. SBA applies a set of editing rules and procedures to those data and links each establishment in the DMI files to the data on that establishment from previous years. If the establishment is part of a multi-establishment company, the establishment is also linked to the rest of the information about the company. The resulting web of information allows SBA to track employment growth in groups of establishments over time and to compare the attributes of large and small businesses.
Over the years the comparison or benchmarking of the data from USEEM with other sources of business data has resulted in the appearance of inconsistencies. While some differences will be found in almost any direct comparison because of different definitions and assumptions underlying the data, new editing procedures and refinements of old ones have been introduced to bring the data reported in USEEM closer to the data reported in other data sources. However, some inconsistencies still exist. The purpose of this project was to determine the accuracy of the information in the DMI files and the validity of the editing procedures used by SBA to construct USEEM from those files. And from that analysis infer which problems might be affecting the quality of the data. The questions previously listed in the summary report were assembled to help structure that analysis.

The process of validating the entire data base was too large an undertaking for a project of this size. Since many of the questions addressed issues in the linking of establishment records to create entire multi-establishment firms, it was decided that this study should focus on large multi-establishment employers.

Preliminary Comparisons with Public Data Sources

To verify a data set requires a second source of information with which to compare the original data. A survey was thought to be the best way of obtaining this second set of data for this
study. However, because of the expense of a survey other possible data sources were explored first.

For use in the preliminary data analysis and the test surveys, the SBA provided a list of 50 firms from the 1986 USEEM that had been identified by Applied Systems Institute (the firm that produced the 1976-86 USEEM files from the DMI files) as being large firms with either a substantial under reporting of branch or subsidiary establishments or with a high likelihood that several establishments were consolidated into one very large establishment.

SEC 10-K Data

The first step was to compare the 1986 USEEM data on those 50 firms with public data sources. The list of 50 was cross checked against the list of firms that filed 10-K reports with the Securities and Exchange Commission. Twenty-one of the firms had reports that could be located and the employment, sales, number of subsidiaries and industry code from their 10-K forms were compared to the data in the USEEM files.

Table 1 shows the comparison for each firm, but several general observations could be made. The industry classifications as shown by the SIC codes tend to be quite close. For 16 of the
<table>
<thead>
<tr>
<th>Firm</th>
<th>Employment USEEM % different from 10K</th>
<th>Sales-Year USEEM Sales closest to 10K Sales</th>
<th># Subsidiaries on 10K and not on USEEM</th>
<th># Subsidiaries on USEEM and not on 10K</th>
<th>Identifiable Foreign Subsidiaries</th>
<th>SIC Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>1985</td>
<td>4</td>
<td>2</td>
<td>No</td>
<td>4-digit</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>1985</td>
<td>42</td>
<td>4</td>
<td>No</td>
<td>3-digit</td>
</tr>
<tr>
<td>3</td>
<td>NA</td>
<td>1986</td>
<td>100+</td>
<td>8</td>
<td>Yes</td>
<td>4-digit</td>
</tr>
<tr>
<td>4</td>
<td>-25</td>
<td>1986</td>
<td>2</td>
<td>2</td>
<td>No</td>
<td>3-digit</td>
</tr>
<tr>
<td>5</td>
<td>-4</td>
<td>1986</td>
<td>--</td>
<td>2</td>
<td>No</td>
<td>4-digit</td>
</tr>
<tr>
<td>6</td>
<td>2/</td>
<td>1986</td>
<td>2</td>
<td>2/</td>
<td>2/</td>
<td>2/</td>
</tr>
<tr>
<td>7</td>
<td>29</td>
<td>1986</td>
<td>2</td>
<td>6</td>
<td>No</td>
<td>3-digit</td>
</tr>
<tr>
<td>8</td>
<td>199</td>
<td>1985</td>
<td>27</td>
<td>1</td>
<td>Yes</td>
<td>3-digit</td>
</tr>
<tr>
<td>9</td>
<td>248</td>
<td>1985</td>
<td>22</td>
<td>1</td>
<td>Yes</td>
<td>3-digit</td>
</tr>
<tr>
<td>10</td>
<td>54</td>
<td>1985</td>
<td>6+</td>
<td>1</td>
<td>Yes</td>
<td>No match</td>
</tr>
<tr>
<td>11</td>
<td>115</td>
<td>1985</td>
<td>56</td>
<td>4</td>
<td>Yes</td>
<td>3-digit</td>
</tr>
<tr>
<td>12</td>
<td>-4</td>
<td>1985</td>
<td>--</td>
<td>--</td>
<td>No</td>
<td>3-digit</td>
</tr>
<tr>
<td>13</td>
<td>-3</td>
<td>1985</td>
<td>--</td>
<td>1</td>
<td>No</td>
<td>4-digit</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>1985</td>
<td>--</td>
<td>--</td>
<td>No</td>
<td>4-digit</td>
</tr>
<tr>
<td>15</td>
<td>-3</td>
<td>1986</td>
<td>3</td>
<td>9</td>
<td>No</td>
<td>No match</td>
</tr>
<tr>
<td>16</td>
<td>NA</td>
<td>1985</td>
<td>--</td>
<td>5</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>17</td>
<td>-9</td>
<td>1986</td>
<td>8</td>
<td>--</td>
<td>No</td>
<td>4-digit</td>
</tr>
<tr>
<td>18</td>
<td>2/</td>
<td>1985</td>
<td>4</td>
<td>2</td>
<td>No</td>
<td>4-digit</td>
</tr>
<tr>
<td>20</td>
<td>2/</td>
<td>1985</td>
<td>5</td>
<td>1</td>
<td>No</td>
<td>4-digit</td>
</tr>
<tr>
<td>21</td>
<td>5/</td>
<td>1985</td>
<td>6</td>
<td>3</td>
<td>No</td>
<td>4-digit</td>
</tr>
</tbody>
</table>

1/USEEM sales 18% smaller than closest possibility on 10K.

2/Listed as a subsidiary on another company's 10K - No separate information available.

3/Corporate sales not available on USEEM and sum of subsidiaries larger than amount listed on 10K.

4/Listed as subsidiary on another company's 10K but sold during 1986. No information on new owner available.

5/Company declared bankruptcy 1/86, information from 1985 10K.
21 firms, the sales numbers matched pretty closely, 69 percent (11 firms) were fiscal 1985 sales and 31 percent (5 firms) were fiscal 1986 sales. In every case where fiscal 1986 sales were available, the firm was not on a calendar year fiscal year and in all but one case the fiscal year ended during the first half of the calendar year. This is a very reasonable finding since the USEEM file is closed at the end of the calendar year before most firms would have calculated calendar year sales. And for the large publicly held companies that are most likely to be reporting to the SEC, it is probably much easier to give Dun & Bradstreet the latest available full fiscal year of consolidated sales information than it is to make an estimate of current year sales.

It is not as easy to generalize about the employment comparisons. The SEC does not specify how employment is to be reported on the 10-K; therefore, the only thing known is that it is an employment count for each company's 1986 fiscal year. About half the firms showed employment within 10 percent of the USEEM number and for five firms it was not possible to compare employment. Of the remaining six firms all but one had significantly more employment reported on USEEM than on the 10-K. Since most of those firms had foreign subsidiaries, it is possible that the 10-K information was only domestic employment and the USEEM total included foreign employment; however, the magnitude

---

13 While company's frequently report their foreign employment in the text or tables of their annual reports, this information was not routinely extracted for this comparison.