ANNUAL MEASURES OF JOB CREATION AND JOB DESTRUCTION CREATED FROM QUARTERLY ES-202 MICRODATA

Joshua C. Pinkston and James R. Spletzer
Bureau of Labor Statistics
2 Massachusetts Avenue NE, Suite 4945, Washington DC 20212

1) Introduction

Following a specific establishment over time in longitudinal microdata is often quite complex, especially through periods of corporate restructuring. Failure to accurately define an establishment as surviving over time breaks a continuous linkage and thus falsely defines both a death and a birth. Although the importance of constructing accurate longitudinal linkages is well known, certain unique issues arise when trying to analyze establishment survival and employment dynamics across a long period of time. In this paper, we highlight the issues involved in extending longitudinal linkage algorithms across more than two consecutive periods of cross-sectional microdata. We illustrate the empirical effects by constructing annual measures of job creation and job destruction using quarterly cross-sectional microdata from the Bureau of Labor Statistics' ES-202 program and the associated longitudinal establishment database.

2) Longitudinal ES-202 Establishment Microdata

All employers subject to state Unemployment Insurance (UI) laws are required to submit quarterly contribution reports detailing their monthly employment and quarterly wages to the State Employment Security Agencies (SESAs). After the microdata are edited and, if necessary, corrected by the State Labor Market Information staff, the states submit these data and other business identification information to the Bureau of Labor Statistics as part of the Covered Employment and Wages (ES-202) program, which is a cooperative endeavor of BLS and the States. The data gathered in the ES-202 program are a comprehensive and accurate source of employment and wages, and provide a virtual census (98%) of employees on nonfarm payrolls. The ES-202 data serve as the sampling frame for BLS establishment surveys. For more information on the ES-202 program, see Farmer and Searson (1995) and U.S. Bureau of Labor Statistics (1997).

The cross-sectional ES-202 microdata are then linked across quarters to create a longitudinal database of establishments. This longitudinal establishment database, referred to as the LDB, will be used by the BLS to generate high quality, high frequency, timely, and historically consistent information regarding not only job creation and job destruction, but also the life cycle of establishments. These statistics will expand our understanding of employment growth, by describing how many establishments are expanding or contracting and by how much these establishments are expanding or contracting. The LDB contains the entire history of quarterly microdata from 1990 through the most recent quarter available. A detailed description of the LDB is given in the April 2001 *Monthly Labor Review* article by Pivetz, Searson, and Spletzer, and details about the LDB record linkage system can be found in Robertson, Huff, Mikkelson, Pivetz, and Winkler (1997).

3) Creating Annual Linkages from Quarterly Microdata: The Technical Issue

As part of the process of linking establishments across quarters, the LDB longitudinal linking algorithm identifies what are termed breakouts and consolidations. The term "breakout" refers to a transition from a single establishment employer to a multi-establishment employer, and the term "consolidation" refers to a transition from a multi-establishment employer to a single establishment employer. Breakouts and consolidations may be actual economic events representing business expansions and contractions, or merely administrative reporting changes due to how an employer with multiple establishments within a state reports its data. Although the BLS and the States continuously work with employers in order to obtain data at the establishment level, some employers with multiple establishments within a state report their total employment and wages in a consolidated manner. Occasionally, an employer reporting consolidated data to a state will disaggregate its data to the worksite level (or, much less frequently, vice-versa).

The record linkage system used to construct the LDB creates flags for establishments involved in a breakout or consolidation. The establishments that are flagged have a one-to-one correspondence with a breakout and consolidation lookup table. For any

given quarter, this lookup table defines the relationships between the establishments that are involved in a 1:N breakout or a N:1 consolidation. Establishments that are involved in a breakout or consolidation often have discontinuous identification numbers (referred to in the remainder of this paper as LDB numbers) across the two quarters. The breakout and consolidation flags, however, alert the analyst that the establishment is a surviving employer involved in a breakout or a consolidation, not an opening or closing as it may appear in the microdata.

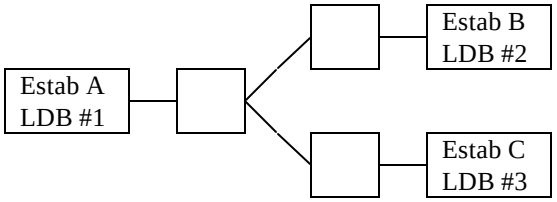
The establishments involved in breakouts and consolidations need to be treated with care when constructing tables of job creation and job destruction or establishment openings and closings for a given quarter. If the breakout and consolidation flags and the associated lookup table are ignored, the longitudinal microdata for a business that undergoes a breakout would appear to be a closing of a single existing establishment and the opening of several new establishments. Similarly, the longitudinal microdata for a business that undergoes a consolidation would appear to be a closing of several existing establishments and the opening of one new establishment.

To treat the breakouts and consolidations correctly, the establishments involved in a breakout need to be collapsed according to the relationships defined in the lookup table. This collapsed unit can then be compared to its single establishment partner in the previous quarter. Similarly, the establishments in the quarter before a consolidation need to be collapsed according to the relationships defined in the lookup table, and this collapsed unit can then be compared to its single establishment partner in the following quarter.

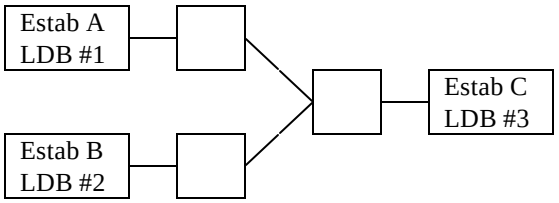
Breakouts and consolidations cause additional problems when trying to compare two points in time that are more than one quarter apart. When the analyst wants to do a comparison from March of one year to March of the following year, information on breakouts and consolidations from all quarters within the year needs to be taken into account in order to understand business continuity and thus avoid spuriously defining openings and closings.

Examples of breakouts and consolidations in the longitudinal ES-202 establishment microdata are illustrated in Figure 1. When describing Figure 1, we find it useful to frame our discussion in terms of an annual comparison from March of one year to March of the next year. In the top of Figure 1, we illustrate the example where establishment A, with LDB number 1 in March of one year, breaks out into two establishments sometime during the year, and these two establishments have LDB numbers 2 and 3 in March of the following year. Roughly 0.74 percent of the establishments in our 1999-2000 California microdata are involved in a breakout. Breakouts are identified by predecessor numbers in the ES-202 microdata, and the LDB record linkage system verifies the breakout as a continuous business by comparing total employment across the two quarters.

Figure 1: Breakout



Consolidation



Breakouts usually refer to an administrative change from a consolidated reporting unit to a reporting of individual establishments. These breakouts most often occur within State specific UI numbers, where only the reporting unit identifiers differ across quarters. However, in our data, just over ten percent of the establishments involved in breakouts change their State specific UI number, which often reflects a business splitting off one of its divisions, and this new establishment is set up in the State UI system as a new legal entity.

Generalizing the example of a breakout in the top of Figure 1, we observe many situations where establishment A in March of one year breaks out into multiple establishments sometime during the year, but only one of these establishments, with a different LDB number, survives into March of the following year. It is important to note how this example differs from the simpler situation where an establishment changes its ownership during the year without incurring a breakout or a consolidation. When an establishment changes ownership, it is allowed to change its State specific UI number. But this change will likely be identified by a State supplied predecessor or successor number or by the probabilistic weighted match in the LDB record linkage system, and as such, the LDB number in the BLS longitudinal establishment database remains

constant through this period of corporate restructuring.

The example in the bottom of Figure 1 describes a consolidation, which usually refers to an administrative reporting change from a reporting of individual establishments to a reporting of a consolidated reporting unit. Consolidations occur much less frequently than breakouts: the incidence rate for breakout is roughly 0.74 percent, whereas the incidence rate for consolidations is only about 0.07 percent.

4) A Description of Two Possible Longitudinal Linking Algorithms

In this section, we describe and compare two methodologies for creating annual tabulations of job creation and job destruction from the quarterly ES-202 establishment microdata. The first is a naïve approach, which takes two quarters of microdata that are one year apart and links them by LDB number without accounting for any breakouts and consolidations that occur within the year. The second is what we view as the correct approach, and uses all information on breakouts and consolidations within the year to assist in defining establishment survival and thus minimizes the number of spurious openings and closings. From a data processing point of view, the correct approach is much more complicated than the naïve approach since it involves merging in the breakout and consolidation lookup table for all relevant quarters within the year and collapsing all flagged establishments into an aggregated employer for the March to March comparison.

Before going to the data, we should describe the theoretical differences between the two longitudinal linking algorithms. We would expect to see more openings and closings with the naïve linking algorithm compared to the correct annual linkages, and we would expect to see more continuous establishments with the correct linking algorithm compared to the naïve annual linkages. In the example of the breakout in the top of Figure 1, the naïve algorithm would define establishment A as a closing and establishments B and C as openings, while the correct algorithm would classify establishment A and the aggregation of establishments B and C as continuous.

5) The Data and Empirical Results

Our goal in the following empirical work is to document the differences between the naïve annual linking algorithm and the correct annual linking algorithm. In the empirical work that follows, we use data from California between the first quarter of 1999 and the first quarter of 2000. Results from other years and states are consistent with those we report here, but are not presented for the sake of brevity.

The empirical comparison of the two linkage methodologies is presented in Table 1. Between the first quarter of 1999 and first quarter of 2000, the California labor market grew from 11,512,734 jobs to 11,895,768 jobs. Using the growth rate defined in Pivetz, Searson, and Spletzer (2001), the number of jobs grew by 3.3 percent. The net employment growth of 383,034 jobs occurred as some establishments expanded, some contracted, and some establishments either opened or closed. Job creation is defined as employment growth contributed by establishments that expanded or opened, and job destruction is defined as the employment decline resulting from establishments that contracted or closed. The sum of job creation and job destruction is the net change in employment.

Using the naïve linkage algorithm, employment in expanding establishments grew by 1,310,823 jobs, and opening establishments were responsible for 1,035,786 new jobs. Employment in contracting establishments declined by 941,415 jobs, and closing establishments accounted for the loss of 1,022,160 jobs. The naïve linkage algorithm puts the job creation rate at 20.0 percent and the job destruction rate at 16.8 percent.

Using the correct linkage algorithm, the job creation rate is 18.7 percent and the job destruction rate is 15.4 percent. The difference between both the job creation rate and the job destruction rate calculated using the two different methodologies is 1.4 percentage points. Using absolute numbers rather than percentages, the naïve linking algorithm relative to the correct linking algorithm inflates annual job creation and annual job destruction by 160,586 jobs each. We see from Table 1 that most of this difference arises from jobs gained from establishment openings or from jobs lost from establishment closings. For example, the naïve linking algorithm says that over one million jobs are lost due to closings, whereas the correct linking algorithm says that only 826,487 jobs are lost due to closings. This difference in the opening and closing statistics is what the discussion above predicted, since the naive linkage algorithm was expected to result in spurious openings and closings.

The establishment counts underlying these job creation and job destruction statistics are also given in Table 1. Using the naïve linking algorithm, there are 675,595 establishments with positive employment in March 1999, and 692,526 establishments with positive employment in March 2000. There are 205,139 establishments (30.0 percent) expanding during the year, and 172,752 establishments (25.3

percent) contracting during the year. There are 106,968 establishments (15.6 percent) opening during the year, and 90,037 establishments (13.2 percent) closing during the year. The difference between the number of establishments opening and closing (16,931 establishments) is the change in the number of establishments between March 1999 and March 2000.

Using the correct linking algorithm, there are 675,241 establishments with positive employment in March 1999, and 688,216 establishments with positive employment in March 2000. These cross-sectional establishment counts differ from those reported for the naïve linking algorithm. In the correct linking algorithm, establishments in March 2000 that are involved in a 1:N breakout since March 1999 are aggregated into an employer specific record for comparison back to the single employer record. Similarly, establishments in March 1999 that are involved in a N:1 consolidation before March 2000 are aggregated into an employer specific record for comparison forward to the single employer record. Comparing the two linkage methods, the difference in the number of establishments in March 2000 (4,310 establishments, as reported in Table 1) reflects the aggregation of establishments involved in breakouts. Similarly, the difference in the number of establishments in March 1999 (354 establishments) reflects the aggregation of establishments involved in consolidations.

Taking this methodologically induced difference one step further, the correct linking algorithm leads to what may appear to be inconsistencies across time in the establishment counts. Specifically, note that the number of March 1999 establishments reported in the March 1999 to March 2000 annual comparison will not equal the number of March 1999 establishments reported in the March 1998 to March 1999 annual comparison. This reflects the establishments involved in breakouts and consolidations being aggregated into an employer specific record only in the year the breakout or consolidation occurs. However, the employment counts will always be identical, not only for a comparison of methodologies, but also across annual comparisons of different years.

The number of establishments classified as openings is 0.7 percentage points smaller in the correct method than in the naïve method (14.95 percent versus 15.64 percent), and the number of establishments classified as closings is 0.1 percentage points smaller in the correct method. The disparity between the 0.7 and the 0.1 statistics reflects the interaction of two factors. First, as noted earlier, the number of breakouts we observe in the data is an order of magnitude larger than the number of consolidations. Second, when considering an 1:N breakout with the naïve method, the number of spurious openings is by definition larger than the number of spurious closings.

There are several additional findings in Table 1 that warrant discussion. Using the correct linkage method, we estimate that the average size of an opening establishment is 8.3 employees (849,513/101,919), and the average size of a closing establishment is 9.3 employees (826,487/88,944). Yet if we look at the difference column, the average size of a spurious opening in the naïve linking algorithm is estimated to be 36.9 employees (186,273/5,049), and the average size of a spurious closing in the naïve linking algorithm is estimated to be 179.0 employees (195,673/1,093). These statistics suggest that although we are only changing the classification of a few thousand spurious openings and closings, these are very large establishments on average and thus we are changing the job creation and job destruction employment statistics by a relatively large amount.

Finally, we would like to make a few remarks about the economic interpretations of the job creation and job destruction statistics in Table 1. These job flow statistics reveal the tremendous amount of churning underlying the annual net employment growth rate of 3.3 percent. The sum of the job creation and job destruction rates, which is 34.1 percent, tells us that more than one in three jobs is either created or destroyed between March 1999 and March 2000. Specifically, 18.7 percent of jobs in March 2000 did not exist one year earlier, and 15.4 percent of jobs in March 1999 do not exist one year later. Furthermore, 15.0 percent of establishments opened and 13.0 percent of establishments closed between March 1999 and March 2000. These statistics demonstrate that there are a sizable number of jobs and businesses that appear and disappear during the relatively short time frame of one year.

The job creation rate of 18.7 percent and the job destruction rate of 15.4 percent reported in Table 1 are somewhat higher than in the relevant literature. Spletzer (2000) reports annual job creation and job destruction rates of 14.6 percent and 13.2 percent, respectively, using data from West Virginia in the early 1990s. The rates for manufacturing reported by Spletzer (2000) are 10.4 percent and 13.7 percent. The annual job creation and job destruction rates reported by Davis, Haltiwanger, and Schuh (1996), using manufacturing data from 1973-1988, are 9.1 percent and 10.3 percent, respectively. Two immediate explanations for the somewhat higher rates we find are state effects and time period effects: it may be possible that California has higher job reallocation than other states, or it may be possible

that the late 1990's and early 2000's have higher job reallocation than earlier years. In future research, we plan to quantify these possible state and year effects.

6) Conclusions

In this paper, we have discussed the construction of annual job creation and job destruction statistics using quarterly establishment level microdata from the Unemployment Insurance (UI) systems of the 50 states plus the District of Columbia. Our discussion and empirical results show that methodology matters. Differences in how the microdata are linked over time result in relatively large effects on the gross job flow statistics.

While this paper concentrates on the methodology necessary for producing annual tabulations from quarterly microdata, the ultimate goal of this project is to produce both an algorithm and a database that allows for longitudinal tabulations at other frequencies such as biennial, triennial, quinquennial, and decennial. The longitudinal ES-202 database resulting from this work will cover nearly all establishments in all industries, and will provide an excellent source of data for research into topics such as employment adjustment, corporate restructuring, and business survival.

Disclaimer

All empirical work in this paper is based on the authors' calculations and is exploratory research meant to motivate discussion about methodology. Any views expressed in this paper are those of the authors and do not necessarily reflect the policies of the BLS or the views of other BLS staff members.

References

Davis, Steven J., John C. Haltiwanger, and Scott Schuh. 1996. *Job Creation and Destruction*. MIT Press.

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." *Proceedings of the Bureau of the Census 1995 Annual Research Conference*, pp. 198-235.

Pivetz, Timothy R., Michael A. Searson, and James R. Spletzer. 2001. "Measuring Job Flows and Establishment Flows With BLS Longitudinal Establishment Microdata." *Monthly Labor Review*, Vol. 124, No. 4, April 2001, pp. 13-20.

Robertson, Kenneth, Larry Huff, Gordon Mikkelson, Timothy Pivetz, and Alice Winkler. 1997. "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.

Spletzer, James R. 2000. "The Contribution of Establishment Births and Deaths to Employment Growth." *Journal of Business and Economic Statistics*, Vol. 18, No. 1, January 2000, pp. 113-126.

U.S. Bureau of Labor Statistics, 1997. *BLS Handbook of Methods, Bulletin #2490*.

Table 1: Annual Estimates

Annual Employment Levels and Flows, California March 1999 – March 2000

Naïve Linkage Method		
Employment	Level	Percent
March 1999	11,512,734	
March 2000	11,895,768	
Change	383,034	3.3
Job Creation		
Total	2,346,609	20.0
Expanding	1,310,823	11.2
Opening	1,035,786	8.8
Job Destruction		
Total	1,963,575	16.8
Contracting	941,415	8.0
Closing	1,022,160	8.7

Correct Linkage Method		
Employment	Level	Percent
March 1999	11,512,734	
March 2000	11,895,768	
Change	383,034	3.3
Job Creation		
Total	2,186,023	18.7
Expanding	1,336,510	11.4
Opening	849,513	7.3
Job Destruction		
Total	1,802,989	15.4
Contracting	976,502	8.3
Closing	826,487	7.1

Difference (Correct – Naïve)		
Employment	Level	Percent
March 1999	0	
March 2000	0	
Change	0	0.0
Job Creation		
Total	-160,586	-1.4
Expanding	25,687	0.2
Opening	-186,273	-1.6
Job Destruction		
Total	-160,586	-1.4
Contracting	35,087	0.3
Closing	-195,673	-1.7

Annual Establishment Levels and Flows, California March 1999 – March 2000

Naïve Linkage Method		
Number Establishments	Level	Percent
March 1999	675,595	
March 2000	692,526	
Change	16,931	2.5
Job Creation		
Total	312,107	45.6
Expanding	205,139	30.0
Opening	106,968	15.6
Job Destruction		
Total	262,789	38.4
Contracting	172,752	25.3
Closing	90,037	13.2

Correct Linkage Method		
Number Establishments	Level	Percent
March 1999	675,241	
March 2000	688,216	
Change	12,975	1.9
Job Creation		
Total	307,430	45.1
Expanding	205,511	30.1
Opening	101,919	15.0
Job Destruction		
Total	262,036	38.4
Contracting	173,092	25.4
Closing	88,944	13.0

Difference (Correct – Naïve)		
Number Establishments	Level	Percent
March 1999	-354	
March 2000	-4,310	
Change	-3,956	-0.6
Job Creation		
Total	-4,677	-0.5
Expanding	372	0.2
Opening	-5,049	-0.7
Job Destruction		
Total	-753	-0.0
Contracting	340	0.1
Closing	-1,093	-0.1