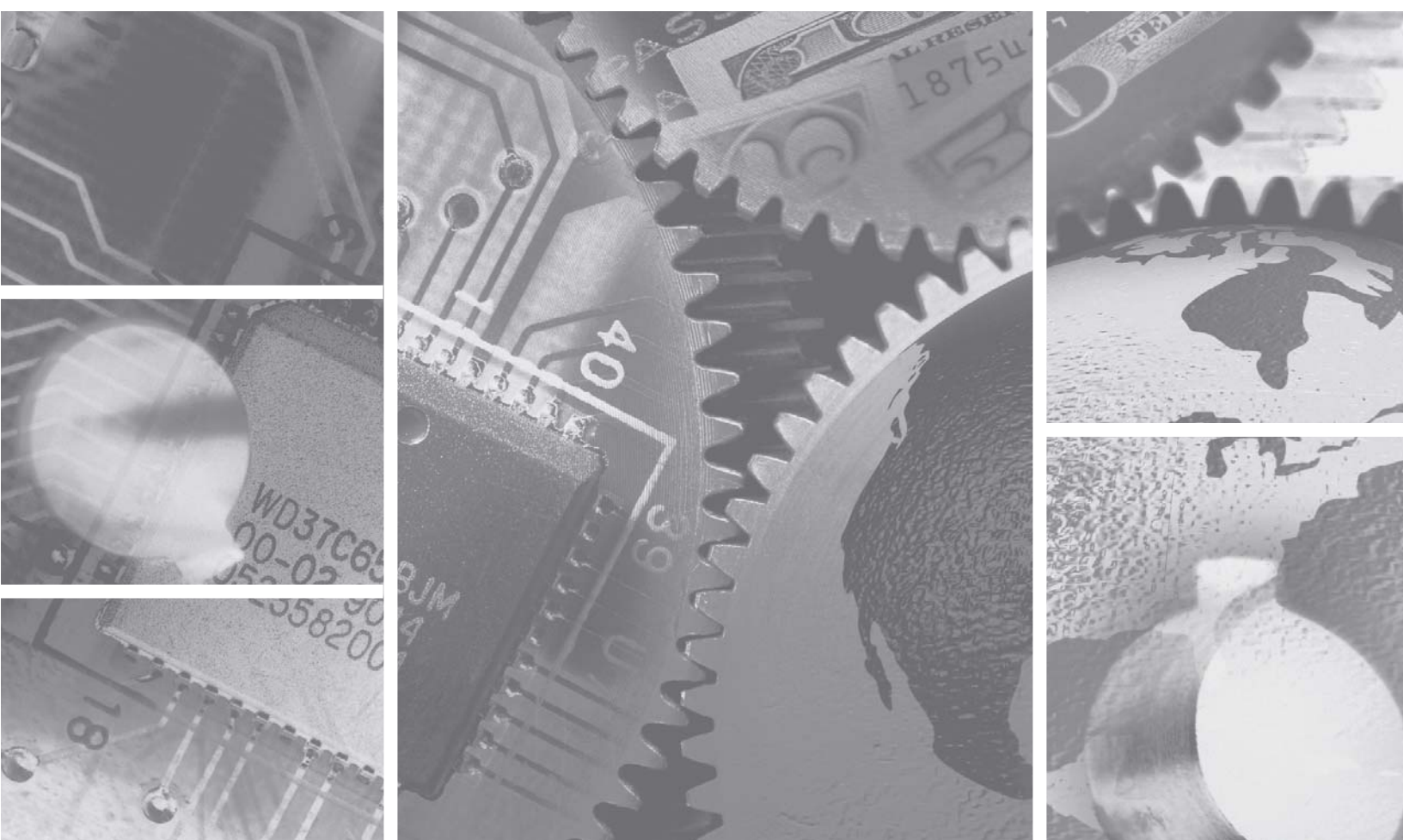BUSINESS DYNAMICS STATISTICS:

An Overview

John Haltiwanger
University of Maryland

Ron Jarmin
U.S. Bureau of the Census

Javier Miranda
U.S. Bureau of the Census

KAUFFMAN
The Foundation of Entrepreneurship

"The fundamental impulse that keeps

the capital engine in motion comes

from the new consumers' goods,

the new methods of production

and transportation, the

new markets …

[The process] incessantly

revolutionizes from within,

incessantly destroying the old one,

incessantly creating a new one.

This process of Creative Destruction

is the essential fact

of capitalism."

– Joseph Schumpeter

TABLE OF CONTENTS

Business Dynamics Statistics— General Explanation

> The population of these businesses is constantly churning—some businesses grow, others decline, and yet others close.

Introduction

The Business Dynamics Statistics (BDS) includes measures of establishment openings and closings, firm startups, job creation and destruction by firm size, age, and industrial sector, and several other statistics on business dynamics. The U.S. economy is comprised of more than 6 million establishments with paid employees. The population of these businesses is constantly churning—some businesses grow, others decline, and yet others close. New businesses constantly replenish this pool. The BDS series provides annual statistics on gross job gains and losses for the entire economy, and by industrial sector and state. These data track changes in employment at the establishment level, and thus provide a picture of the dynamics underlying aggregate net employment growth.

Sources of Data

The Business Dynamics Statistics is a product of the U.S. Census Bureau and was developed by the Center for Economic Studies. The BDS data are compiled from the Longitudinal Business Database (LBD) (Jarmin and Miranda, 2002). The LBD is a longitudinal database of business establishments and firms covering the years between 1976 and 2005 (as additional years of the LBD are available, the BDS will be updated).

The LBD is constructed by linking annual snapshot files from the Census Bureau's Business Register (BR)[1] to provide a longitudinal history for each establishment. The linkage process makes use of numeric establishment identifiers as well as probabilistic name and address matching. The linkage process allows the tracking of net employment changes at the establishment level, which, in turn, allows the estimation of jobs gained at opening and expanding establishments, and jobs lost at closing and contracting establishments.

The LBD originally was conceived and constructed to be a research dataset. It has been used in numerous studies published in leading scholarly journals. It also has seen increased use as a source of special tabulations. The growing demand for tabulations from the LBD is why the Census Bureau has developed the BDS.

1. The U.S. Census Bureau has maintained a general-purpose business register since 1972. The Business Register (formerly known as the Standard Statistical Establishment List, or SSEL) describes U.S. business establishments and companies. The Business Register is continuously updated with administrative data from other federal agencies, as well as data collected by the Census Bureau. Information for single establishments and administrative units of multi-establishment companies (EINs) is updated based on payroll tax records from the Internal Revenue Service. Information for establishments of multi-unit companies is updated annually based on responses to the Company Organization Survey. Other routine sources of update information include Census Bureau current surveys (e.g., Annual Survey of Manufacturers) and the economic census.

Coverage, Scope, and Key Definitions

The BDS' coverage and scope are based on the Census Bureau's County Business Patterns (CBP) program. The CBP program has been producing annual statistics on measures of economic activity since 1964. The CBP serves a wide constituency of users, including researchers, policymakers, and the business community. It is used by the research community to study the economic activity of small areas and to analyze economic change over time; and by program areas as a benchmark for statistical series, surveys, and databases between economic censuses. Businesses use the data for analyzing market potential, measuring the effectiveness of sales and advertising programs, setting sales quotas, and developing budgets. Government agencies use the data for administration and planning.

Businesses use the (CBP) data for analyzing market potential, measuring the effectiveness of sales and advertising programs, setting sales quotas, and developing budgets.

Table 1. BDS Industrial Scope and Coverage, and Definitions of Key Data Items

• Sectors (SIC-Based)	agricultural services, forestry, and fishing
	mining
	construction
	manufacturing
	transportation and public utilities
	wholesale trade
	retail trade
	finance, insurance, and real estate
	services
Excluded	Self-employed, domestic service workers, railroad employees, agricultural production workers, most government employees, and employees on ocean-borne vessels or in foreign countries.
• Establishments	The CBP series excludes governmental establishments except for liquor stores (SIC 592), wholesale liquor establishments (SIC 518), depository institutions (SIC 60), federal and federally sponsored credit agencies (SIC 611), and hospitals (SIC 806).
• Employment	Full- and part-time employees for pay period that includes March 12 of each year. Included are employees on paid sick leave, holidays, and vacations. Not included are proprietors and partners of unincorporated businesses.
• Payroll	Total payroll includes all forms of compensation, such as salaries, wages, reported tips, commissions, bonuses, vacation allowances, sick-leave pay, employee contributions to qualified pension plans, and the value of taxable fringe benefits.

Table 1 presents a detailed listing of the CBP in-scope rules. CBP covers most of the country's economic activity. The only major exclusions are self-employed individuals, employees of private households, railroad employees, agricultural production employees, and most government employees. This represents the practical totality of the private, non-agricultural sector of the economy.

Establishments and Firms

Establishments are used in tabulating the BDS statistics. An establishment is a fixed physical location where economic activity occurs. A firm may have one establishment (a single-unit establishment) or many establishments (a multi-unit firm). Firms are defined at the enterprise level such that all establishments under the operational control of the enterprise are considered part of the firm. Firm-level data are compiled based on an aggregation of establishments under common ownership by a corporate parent using Census Bureau company identification numbers. The firm-level aggregation, which is consistent with the role of corporations as the economic decision makers, is used for the measurement of the LBD data elements by size class and age class.

Concepts and Methodology

The BDS data measure the net change in employment at the establishment level. These changes come about in one of four ways. A net increase in employment can come from either *opening establishments* or *expanding establishments*. A net decrease in employment can come from either *closing establishments* or *contracting establishments*.

Gross job gains include the sum of all jobs added at either opening or expanding establishments. *Gross job losses* include the sum of all jobs lost in either closing or contracting establishments. The net change in employment is the difference between gross job gains and gross job losses.

The formal definitions of employment changes are:

Job Creation (JC): Job creation is the sum of all employment gains from expanding establishments from year *t-1* to year *t*, including establishment startups. Note that the contribution of firm births can be measured by using the job creation from establishments with firm age equal to zero.

Job Destruction (JD): Job destruction is the sum of all employment losses from contracting establishments from year *t-1* to year *t*, including establishments shutting down.

Some simple identities are useful to note to interpret and use these statistics. Let E_{it} be employment in year *t* for establishment *i*.

Define establishment-level employment growth as follows:

$$g_{it} = (E_{it} - E_{it-1})/X_{it}$$

where

$$X_{it} = .5*(E_{it} + E_{it-1})$$

This growth-rate measure has become standard in analysis of establishment and firm dynamics because it shares some useful properties of log differences, but also accommodates entry and exit (see Davis, Haltiwanger, and Schuh [1996], and Tornquist, Vartia, and Vartia [1985]). The above definitions of JC and JD for establishments classified in group s (e.g., a firm-size, firm-age category) are given by:

$$JC_{st} = \sum_{\substack{i \in s, \\ g_{it} \geq 0}} (E_{it} - E_{it-1}) \qquad JD_{st} = \sum_{\substack{i \in s, \\ g_{it} < 0}} |(E_{it} - E_{it-1})|$$

The net change in employment for establishments in group *s* satisfies the following identity:

$$NET_{st} = \sum_{i \in s} (E_{it} - E_{it-1}) = JC_{st} - JD_{st}$$

For growth rates, the analogous relationships are given by:

$$JCR_{st} = \sum_{\substack{i \in s, \\ g_{it} \geq 0}} (X_{it}/X_{st})g_{it} = JC_{st}/X_{st}$$

$$JDR_{st} = \sum_{\substack{i \in s, \\ g_{it} \geq 0}} (X_{it}/X_{st})|g_{it}| = JD_{st}/X_{st}$$

$$NETR_{st} = \sum_{i \in s} (X_{it}/X_{st})g_{it} = NET_{st}/X_{st} = (JC_{st} - JD_{st})/X_{st}$$

where

$$X_{st} = \sum_{i \in s} X_{it}$$

The latter variable X_{st} denotes the sum of average employment over a consecutive two-year period and, as is clear from the above, it is simple to convert the changes to rates by dividing the relevant measures by this variable. Note that, in general, the variable X_{st} for a particular classification is not equal to the simple average of the employment variable using the current and prior year, because establishments are assigned the characteristics of the firm that owns the establishment in *t* and this may have changed from year *t-1* to year *t*.

The employment measure used for the tabulations is the number of employees at the establishment for the payroll period that includes March 12. As such, all growth rates are based on March-to-March changes, and the tabulations for a given year are the changes from the prior year to the current year. An establishment opening or entrant is an establishment with positive employment in the current

This growth-rate measure has become standard in analysis of establishment and firm dynamics because it shares some useful properties of log differences, but also accommodates entry and exit.

The vast majority of establishment openings are true greenfield entrants. Similarly, the vast majority of establishment closings are true establishment exits (i.e., operations ceased at this physical location).

year and zero employment in the prior year. An establishment closing or exit is an establishment with zero employment in the current year and positive employment in the prior year. The vast majority of establishment openings are true greenfield entrants. Similarly, the vast majority of establishment closings are true establishment exits (i.e., operations ceased at this physical location). However, there are a small number of establishments that temporarily shut down (i.e., have a year with zero employment), and these are counted in the establishment openings and closings.

In the released series, the job flow measures are provided in terms of both level changes (e.g., the number of jobs) and rates using the denominator as described above to convert level changes to rates. In addition, the number of establishments in each of the categories of change (openings, closings, continuers) and the classifications (e.g., firm size, firm age, etc.) is provided, permitting tracking of the gross and net flows of the number of establishments. The decomposition into openings, closings, and continuers permits decomposing gross job creation into the component from continuing establishments that are expanding and establishment openings, and decomposing gross job destruction into the component from continuing establishments that are contracting and establishment closings.

It is critical to emphasize that the BDS contains measures of net and gross flows of establishments and jobs at the establishment level. All establishments are, however, linked to their parent firm so that the net and gross flows of establishments and jobs can be categorized by the characteristics of the parent firm. In particular, establishments are classified by both the size of the parent firm and age of the parent firm as defined below. This enables quantifying the contribution of firms by firm size and firm age in terms of establishment and job net and gross flows. For example, and of particular interest, the contribution of firm startups to the net and gross flows of establishments and jobs can be ascertained by using the tabulations of firm age zero. As described in detail below, establishments are assigned a firm age based upon the age of the parent firm. The age of the parent firm is based on the age of the oldest establishment in the firm. A firm age of zero represents a firm where all establishments in the firm are entrants in that year—hence, it is a new firm. By construction, tabulations of firm age zero represent establishment entrants that are part of a new firm. Most new firms are single-unit firms.

Firm Size and Firm Age Measurement

The BDS describes job flow and entry/exit patterns for establishments in the United States. When describing these patterns by size and age, we classify establishments according to the size and age of the firm with which they are associated. We use two measures of firm size: initial year and actual year. Actual firm size is defined as

the average of year *t-1* and year *t* employment. Initial firm size is defined for any given consecutive two-year period as the size at year *t-1*, except in cases when year *t-1* employment is equal to zero, in which case initial size is year *t* employment. Initial firm size categories are identical to those for firm size.

Both establishment and firm age are computed from the data (see Jarmin, 2002). Birth year is defined as the year an establishment first reports positive employment in the LBD. Establishment age is computed by taking the difference between the current year of operation and the birth year. Given that the LBD series starts in 1976, observed age is by construction left censored at 1975.

Firm age is computed for all firms in the LBD from the age of the establishments belonging to that particular firm. A firm is assigned an initial age by determining the age of the oldest establishment that belongs to the firm at time of birth. Firm age accumulates with every additional year after that. Note that mergers and acquisitions and divestitures could lead to abrupt changes in firm age purely from establishment composition issues if we defined firm age in each year using age of the oldest establishment owned in that year.

Comparability with Other Data

Although the BDS is based on the same basic source data as the Census Bureau CBP and Statistics of U.S. Business (SUSB) programs, differences in how the source data are processed lead to differences in the published statistics. The LBD program involves many edits and updates done for longitudinal consistency that are not done for the CBP and SUSB. Since the BDS is sourced from the LBD, these differences are carried through to the BDS tabulations.

Nevertheless, the BDS tracks the CBP and SUSB quite closely. Figure 1 compares total employment in the CBP and LBD series. It is clear that the two series track each other closely. Both the CBP and the LBD use the same underlying data; however, each series imposes its own set of restrictions and edits on the universe of active establishments. One difference is that, after being released, the CBP is not further edited upon subsequent edits to the BR. The LBD series also imposes its own unique set of restrictions and edits, including the use of name and address matching to eliminate duplicates, and the use of algorithms that identify and clean up incorrect data entries in the employment series. Despite these processing differences, the quantitative differences are not large and range within a band of 1.9 million workers, with the LBD

> **Although the BDS is based on the same basic source data as the Census Bureau CBP and Statistics of U.S. Business programs, differences in how the source data are processed lead to differences in the published statistics.**

Figure 1
CBP vs. BDS Employment by Year

typically falling below the CBP numbers, particularly in the second half of the series (due to the unduplication processing for the LBD).

Figure 1 presents results from comparing the CBP versus BDS net employment growth rate series for the in-scope economy by year. As before, the two series track each other reasonably well. The biggest difference, a two-point difference, is found at the beginning of the series. The discrepancy can be tracked mostly to differences in education and mining across the two series and due to the processing differences discussed above.

Figure 2 compares the number of establishment births and deaths in the LBD and SUSB series. The alternative birth and death series track each other well. However, the LBD shows lower numbers, particularly for births during census years (those ending in 2 and 7). The smoother LBD is the result of a series of algorithms designed to identify and retime incorrectly timed births and deaths that come about from census processing activities (see Miranda, 2006). The reassignment of births and deaths result in a smoothing of the series for those years. Note that the retiming algorithms have a clear impact on the birth series but have a weaker impact on the death series. This stems from an asymmetry in the number of incorrect births and deaths (the census does a better job at identifying deaths than births), as well as the relative difficulty in reassigning deaths for multi-unit firms.

Figure 2
Establishment Entry and Exit, Employment Weighted: BDS vs. SUSB by Year
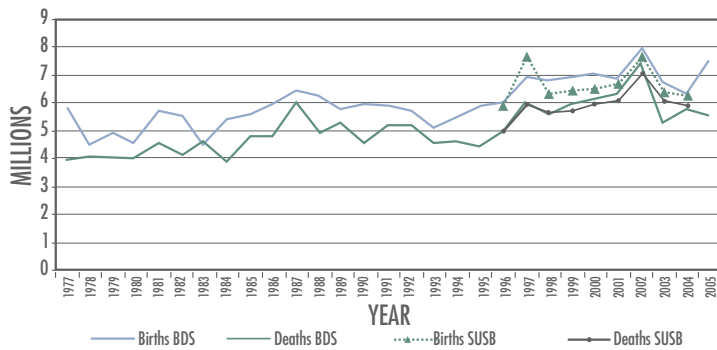

Reliability of the Data

Since the data series on Business Dynamics Statistics is based on administrative rather than sample data, there are no issues related to sampling error. Nonsampling error, however, still exists. Nonsampling errors can occur for many reasons, such as the employer submitting corrected employment data after the end of the year, as well as late filers. Other sources of error include typographical errors made by businesses when providing information. Such errors, however, are likely to be distributed randomly throughout the dataset.

Changes in administrative data sometimes create complications for the linkage process. The LBD addresses these issues in detail to avoid overstating openings and closings, while understating expansions and contractions (Jarmin and Miranda, 2002).

The BDS data series are subject to periodic minor changes based on corrections in LBD records due to updates coming from new BR files (e.g., late filers are no longer considered deaths).

Since the data series on Business Dynamics Statistics is based on administrative rather than sample data, there are no issues related to sampling error.

KAUFFMAN
The Foundation of Entrepreneurship