

## **A Profile of US Data Sources on Entrepreneurship**

Richard Clayton and Jim Spletzer

US Bureau of Labor Statistics

November 29, 2006

*The views expressed in this paper are solely those of the authors and do not necessarily reflect the official positions or policies of the US Bureau of Labor Statistics or the views of other staff members.*

**Introduction:** This paper is designed to list existing sources of economic information on entrepreneurship in the United States. It also attempts to profile their statistical and economic characteristics. This is a work in progress. While there have been several such lists developed previously, the definition and understanding of entrepreneurship continues to evolve, much like the "creative destruction" processes they try to measure. As the definition, or the interpretation of that definition evolves, so must the data sources and the application or interpretation of those sources.

**The Role of Entrepreneurship:** The desire to understand the business formation and business life cycle often is traced back to Schumpeter. Whether understanding creative destruction or other approaches, the importance of promoting entrepreneurship is rooted in the premise that creation of new businesses, new products and service is absolutely critical to the stability and vibrancy of an economy.

An economy cannot survive without entrepreneurship: an economy cannot survive without new businesses replacing dying businesses. As business owners retire or die, many businesses end along with their flow of products and services, the employment status of their staff and that thread in the socio-economic fabric is withdrawn. Without replacements of those threads by new business start-ups or replacements, the fabric deteriorates.

Further, a truly dynamic economy is also characterized by the development of new products and services including new, more efficient methods, where the new methods tend to drive out older, less efficient methods and the businesses that use them. In this way, new methods and more efficient processes replace or drive out inefficient and less desirable products.

Entrepreneurs are broadly defined as those persons or groups that start new businesses and those who shape and reshape their businesses and production processes continually. Thus, policymakers need to pay attention to the role, level and satisfaction of entrepreneurs. To do this effectively, this must be measured and tracked over time.

**Definition of Entrepreneurship:** In developing measure of entrepreneurship, we need a definition, or more likely, a series of definitions, or perhaps a series of measurable factors that collectively attempt to quantitatively capture this economic and social phenomenon. The many writers and theorists incorporate different factors into their definition of

entrepreneurship, including risk taking, innovation, gap-fillers, creative imitators, coordinators and organizers.

Under the innovator definition, is a person building a new fast-food restaurant an entrepreneur? Is a new corner gas station a creative-imitator or merely an imitator? Is this owner a gap-filler for those in the search for gasoline? How can an “entrepreneur” build a new gas station within sight of one or more gas stations and claim the “innovator” or “risk taker” or “gap filler” mantle? Which features and/or factors are we trying to measure? Which of these are measurable? What combination of measures suits the needs of users?

**Uses and Users:** The range of users and uses for relevant data is wide and diverse. Increasingly, users demand timely and usable data at more detailed levels of geography and industry. Users have grown more sophisticated and are uncomfortable with large scale aggregates that do not reflect their precise area of interest.

Users for these data are primarily engaged in two functions:

- 1) understanding of the current and historical level of economic activity
- 2) attempts to influence the level of activity in some manner through some type of intervention to change policies, access to capital, assessing regulatory barriers or other constraints and providing incentives or various sorts.

Therefore, the natural demand is for sufficient data to understand and plan interventions and to monitor the success of these interventions. Changes to tax policy, exemption from selected bureaucratic requirements, etc., have goals of encouraging a targeted area, industry or demographic group. Examples of governmental or bureaucratic requirements are exemptions from certain environmental or labor relations requirements and even economic survey reporting requirements. These are designed to reduce undue costs and burden that might swamp a small business. Those agencies and groups making changes in policies want to see whether the changes had the intended effect. With reasonable timely data, pilot programs could be redesigned if not successful, or expanded if sufficiently successful. Therefore, timely data sources with high frequency, geographic and industry detail and high reliability are preferred.

**Users:**

In the U.S., there is a wide range of users at all levels.

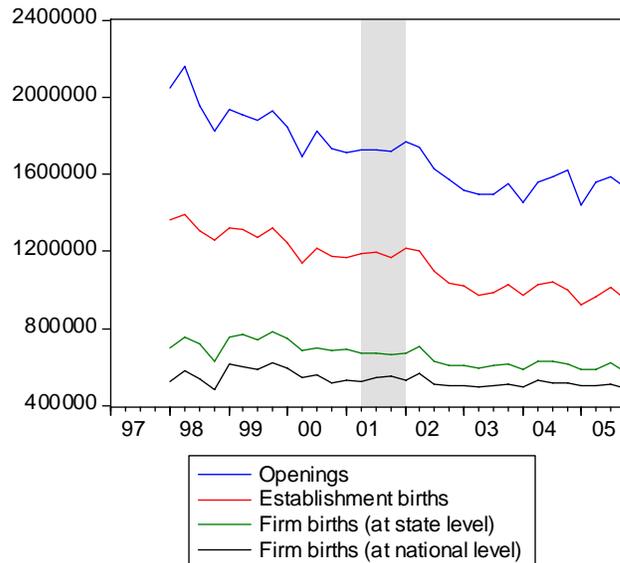
At the federal level, the Small Business Administration is the official advocate for small and new businesses. SBA measures and tracks the situation of small businesses, and also advocates for policy and law changes. At the state and local level, governments seek involvement in the economic development of their jurisdictions. The State of Florida collects its efforts under the banner of “Enterprise Florida”, a design to attract businesses to the state and to provide all of the information and assistance resources in one location to make starting businesses as easy as possible.

At the local level, for example, Fairfax County in northern Virginia outside Washington, D.C. spends \$6.4 million each year employing 34 full-time staff on economic development, and enticements to businesses to locate there. The Fairfax County Economic Development Authority has hosted entrepreneurship fairs, established incubators and science parks. It also has offices in 5 countries seeking businesses that may invest in US operations or expand locations to the US, for the purpose of influencing their decision towards this county. All of this effort and expense is designed to support the ongoing business environment (e.g. jobs, incomes and tax receipts) and living conditions for its citizens.

The many users, each with legitimate interests will want measures that enable them to interpret economic change in their own context, such as by industry or geography. The two charts that follow illustrate these points.

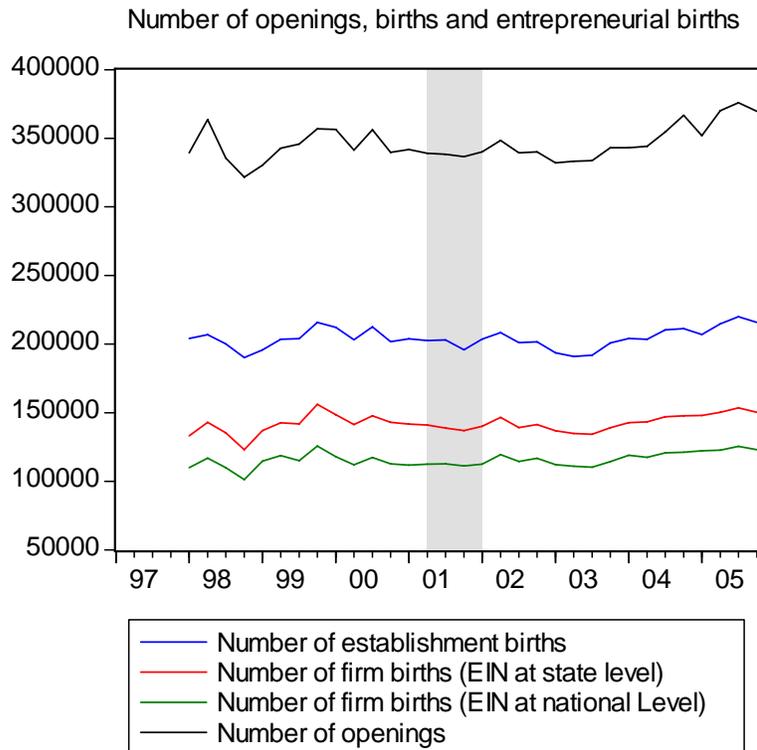
Chart 1.

Jobs from openings, births and entrepreneurial births, seasonally adjusted



Research data from the BLS QCEW program shows varying levels of entrepreneurship based on four different definitions. Openings, including births and seasonal re-openings, has the highest levels. Every quarter, roughly 160,000 or more jobs are created by establishments that opened that quarter. Lower levels of activity are shown for establishment births, lower still for firm level births when constrained to new EINs within each state, and the lowest for firm births measured by new EINs anywhere in the US. These different measures move consistently, but at very different levels. Two things are apparent. First, different users at different levels will desire a measure that captures their particular need. Second, international measures should be carefully considered so as to be as comparable as possible. This, of course, is why this OECD group was established. At the national level, it may be satisfactory to provide measures of new firms, whereas, local users may want establishment level measures that avoid measurement issues with multi-unit businesses that span geographies.

Chart 2:



It is also quickly apparent that a single measure does not answer all questions. While the employment from business births seems to be declining (Chart 1) potentially a cause for concern, the number of business births (Chart 2) is increasing, a more heartening measure. One interpretation is that the average size of a birth is declining, perhaps from increasing productivity. Another is that most new businesses are starting in industries that traditionally require fewer employees. Thus, these same measures are needed by industry.

**Measures of Size class dynamics:** There is considerable literature on the contribution of small and new businesses to economic growth, usually starting with Birch. His work spawned a range of investigations into the proper methodology for measuring this phenomenon.

BLS began publishing job creation and destruction measures by size class in December 2005. This followed a lengthy and detailed review of alternative methods for measuring the contribution of each size class to overall change. While the more commonly used methodologies were considered (base, mean and end sizing), BLS choose what we considered a more balanced and symmetrical approach proposed by Per Davidsson. Under this method, a unit's employment change is spread across the size classes through which it moved from the base to end period of time.

This group may wish to consider this aspect of longitudinal size class methodology as well as the cross sectional definitions.

In addition, two other factors should be included that aid in the utility of the resulting data for current assessment and policy evaluation purposes:

- Frequency: how often data are collected
- Timeliness: lag between the reference period and the release of the data.

**Profile of U.S. Data Sources:** The United States already has a variety of data sources covering a variety of components of the entrepreneurship. These are summarized in the table in attachment 1. This table is in progress and will be expanded to include more specific information on formulas and other information as needed to support the work of this OECD effort.

These sources are derived from both private and public sector sources, households and businesses, administrative data and surveys, and cross-sectional and longitudinal.

These varied sources cover a range of data interests from business births and deaths to the self employed and corporate structures.

**Public Sector Business Data Sources:** At least four federal U.S. government agencies contribute to this field, including the Small Business Administration, the Bureau of Labor Statistics, the Census, and the Bureau of Economic Analysis.

The BLS *Quarterly Census of Employment and Wages* (QCEW/ES-202) program provides likely the most current and comprehensive measures of employment by size, detailed industry and geography, covering all 1200 NAICS industries for each of the 3600 US counties, MSA, states and the nation. Published quarterly and only 7 months after the end of each quarter, the QCEW's combination of frequency, timeliness and detail is unmatched.

In addition, the QCEW provides the basis of the BLS *Business Employment Dynamics* program. The QCEW cross-sectional microdata are linked to provide longitudinal files from 1990 through, at this writing, first quarter 2006. Measuring job creation and destruction, expansions and contractions, opening and closings, the BED are gaining in popularity. These data by size of firm, using a unique allocation methodology were first released in December 2005. The BED is being expanded rapidly to include births and deaths at the firm and establishment levels

The QCEW data also are a key input to the *Longitudinal Employer Household Dynamics (LEHD) program*. At the U.S. Census Bureau, the QCEW establishment files are linked with various sources of individual worker data to provide a range of similar measures with demographic breakdowns by age and gender of the workers.

At the Census Bureau, the periodic economic census and its supporting inputs provides the basis for the *Business Register*, the *Longitudinal Business Database* and the

upcoming *Integrated Longitudinal Business Database (ILBD)*. This latter project seeks to merge business data with that for self-employed. Also, as mentioned, the *LEHD* program links worker demographics with establishment level data to provide new data.

Also, based on the Census business register and the 5-year economic census, the Small Business Administration (SBA) supports the collection of business owner demographics through the *Survey of Business Owners (SBO)* and the *Business Information Tracking Series (BITS)*. The SBO collects information every five years on sex, ethnicity, educational attainment and other

As can be seen, these BLS and Census sources are based on the very comprehensive business lists derived from mandatory state tax reporting (BLS) and federal tax reporting (Census, SBA) again pointing out the lack of survey-based measures that can provide relevant measures.

Also every five years, the Federal Reserve conducts the *Survey of Small Business Finances* for business with less than 500 employees designed to assess the availability of credit to small businesses with an emphasis on minority-owned businesses.

#### **Public Sector Household Data Sources:**

The Bureau of Labor Statistics and the Census Bureau provide data in this area. The BLS *Current Population Survey (CPS)* captures the entrance into entrepreneurship through a monthly household survey. Household members not owning a business in one month will report their ownership in a subsequent month, whether as a sole proprietorship or a business with employees. This source can provide demographics of the business owner.

Also at BLS, the *National Longitudinal Survey (NLS)* provides very detailed survey data for a cohort of the youngest baby boomers. The NLS79 is a survey of 9,964 men and women who were ages 14-22 when first interviewed in 1979, and aged 34-48 when interviewed most recently in 2004. These individuals have been interviewed annually from 1979 to 1994 and biennially since 1994. The NLS data have been widely used to study entrepreneurs (see Fairly, Russman, Schiller).

The *American Community Survey (ACS)* provides cross-sectional data with a longer lag. Also, from the Census Bureau, the decennial census provides incredibly rich demographic data, covering detailed geography each decade.

The Internal Revenue Service compiles federal tax return information under the *Statistics of Income* program illustrating the status of the self-employed and small partnerships.

**Private Sources:** Within the private sector, the foremost sponsor is the Ewing Marion Kauffman Foundation. The largest private survey source is the *Panel Survey of Entrepreneurial Dynamics*, a nationally representative database that includes the characteristics of the population trying to start a new business and characteristics start-up attempts.

The Kauffman Foundation also develops and publishes national and state *Index of Entrepreneurial Activity*, drawn from the monthly BLS Current Population Survey. It also sponsors studies in this field, collects relevant information, supports this field within higher education and other programs.

Also in the private sector is the Dun and Bradstreet *Market Identifier File*. Essentially a private sector business register similar to the BLS QCEW/BED and the Census BR, the Duns file is commercially available for a fee.

It is noticeable that few of the regular economic surveys are listed here. This could be due to several factors, including size class cutoffs in sampling to reduce reporting burden, and also to sample limitations that might allow development of estimates by size and/or age. Also, these regular surveys were historically designed to capture large macroeconomic movements accurately with some degree of industry and geographic detail allowed by sample size and costs. Creating further layers of estimates by size and/or age may not have been considered or deemed cost-effective or sufficiently accurate.

**Strengths and Weaknesses:** Each of these programs and data sources has strengths and weaknesses. Few are current while most have long lags between collections and lags in publication rendering them less than optimal for policy analysis and economic development decisions. Some lack demographics, some are cross-sectional and lack the longitudinal features that allow for short and long term survival studies. However, as a group, this lengthy list has provided answers to most of the questions about the entrepreneurship activity in the US.

**Analysis of Existing Data Sources:** The existing range of data sources covers a wide range – comprehensive business data and a series of longitudinal household-based sources provides insight and richness for studying entrepreneurship.

Most existing business-based sources are from tax-based business registers (BLS and Census) and their derived longitudinal databases. Additional periodic supplements from SBA provide additional detail on the demographics of business owners. These provide excellent geographical and industrial coverage and detail. Some, however, publish data with a substantial lag.

Local and national users of the data primarily seek such data to both understand and to promote entrepreneurship in general or for selected industries or groups. They cannot identify barriers to entry to women as entrepreneurs unless they can determine that women are, or are not, starting and leading businesses at different rates than men. Similarly, those policy makers changing policies, laws or otherwise encouraging selected industries or groups want to monitor the success of their efforts. This calls for frequent profiles with a short time lag.

**Proposed Next Steps:** The process towards achieving the goals of the OECD charter for this group could consider the following:

- A compilation of the needs of data users at all levels: locally, nationally and internationally
- Profiles of existing sources, including frequency and timeliness
- Profiling the existing sources versus the list provided in Davis of Possible Entrepreneurship Indicators
- Comparison of the existing sources for suitability for users
- A statement on data gaps
- A plan for addressing gaps, such as those addressable
  - o through existing sources
  - o through new collections
  - o through a combination of existing and new sources

**Conclusion:** This preliminary draft paper has attempted to profile the existing data sources on entrepreneurship covering the United States. While there are many sources covering a range of aspects, there are opportunities for both expansion of data and for studying the differences among the sources. The work of the OECD group will be important as efforts move forward at national, state and local levels to understand and further the role of entrepreneurship as an economic driving force.

**Acknowledgement:** The authors would like to thank Carol Leming and David Talan of the Business Employment Dynamics program for their assistance in developing and completing the profile of data sources.

**References:**

- Butani, Shail J, Clayton, Richard L., Kapani, Vinod, Spletzer, James R., Talan, David M., Werking Jr., George S., "Business Employment Dynamics: Tabulations by Employer Size," *Monthly Labor Review*, February 2006, 3-22.
- Davidsson, Per "Methodological concerns in the estimation of job creation in different firm size classes." Working Paper, Jönköping International Business School, 1996.
- Davidsson, P., L. Lindmark & C. Olofsson, 1995c, "Small Firms, Business Dynamics and Differential Development of Economic Well-being. *Small Business Economics*," 7, 301-315.
- Davidsson, Per, Leif Lindmark and Christer Olofsson, 1998, "The Extent of Over-estimation of Small Firm Job Creation," *Small Business Economics*, 11(1), 87-100.
- Knaup, Amy E., "Survival and longevity in the Business Employment Dynamics Data," *Monthly Labor Review*, May 2005, 50-56.
- Sadeghi, Akbar, "Birth and Death of Business Establishments: Decomposition of Openings and Closings in Business Employment Dynamics Data Series," Western Economic Association International, July 2005.
- Spletzer, James R., R. Jason Faberman, Akbar Sadeghi, David M. Talan, and Richard L. Clayton, "Business Employment Dynamics: new data on gross job gains and losses," *Monthly Labor Review*, April 2004, 29-42.

- Spletzer, James. "The Contribution of Establishment Births and Deaths to Employment Growth." *Journal of Business and Economic Statistics*, 18(1), January 2000, pp. 113-126.
- Vale, Steve International Data on Business Start ups: Factors Affecting Comparability, Nineteenth International Roundtable of Business Survey Frames, 2006.