Understanding Business Formation and Growth:
The Potential of Linked Employee-Business Data in Entrepreneurial Research

By Chris Goetz, Henry Hyatt, Erika McEntarfer, and Kristin Sandusky

U.S. Census Bureau, Center for Economic Studies and LEHD Program

Entrepreneurship plays a vital role in the U.S. economy. Newly available data on young and small businesses such as the Business Employment Dynamics (BED) and Business Dynamics Statistics (BDS) have enhanced our understanding of the dynamics of young businesses and their important contribution to job creation and economic growth.[1] Yet there remain notable data limitations and measurement challenges in the study of entrepreneurship. In particular, we understand little about the career path of entrepreneurs themselves, the industry experience of startup teams, or the types of jobs created by startups or who is hired into them. The purpose of this paper is to describe the potential for linked employer-employee data to fill some of these data gaps and inform entrepreneurial research and policy.

Linked-employer employee microdata have clear potential to answer many questions of interest in entrepreneurial research. However, approvals for confidential microdata research are lengthy and difficult, putting the data out of reach for many analysts. Because of this, our focus here is on public use data products Census can generate for entrepreneurial research and policy. Specifically, in this paper we describe data products just released, in development, and in the planning stages at the Longitudinal Employer-Household Dynamics (LEHD) program data at the U.S. Census Bureau.[2] Over the last two years, the LEHD program has integrated BDS data on firm age. This allows, for the first time, a link between young firms and their workforce. Census is now developing data products from the linked employer-employee data to produce new statistics on the US economy.

The first new product is Quarterly Workforce Indicators (QWI) by firm age and size, newly released in 2012. The QWI provide hires, separations, turnover, and wages by detailed industry, geography and worker demographics. Prior to the release of these data, we knew little about the characteristics of the jobs generated by young businesses or the demographics of their workforce. The QWI show that startups have a high rate of worker churn relative to established businesses and that startups generally offer lower wages relative to established firms in the same industry (Haltiwanger et al., 2012). Young workers are considerably more likely to work at young businesses than older workers, and have been

[1] See, among others, Dalton et al. (2011), and Haltiwanger, Jarmin, and Miranda (2013).
[2] The LEHD program at Census collects wage data from state Unemployment Insurance (UI) programs, shared with Census via the Local Employment Dynamics (LED) state-federal partnership, linking these job records with firm data to develop a national longitudinal frame of jobs data. As of this writing, coverage is fully national, with the last remaining state submitting data this year.

increasing as a share of employment in startups even while youth employment overall has fallen (Haltiwanger et al., 2013). Employment at startups has also become more diverse in the last decade, with white non-Hispanics constituting a declining share of employment at new businesses.

Another new data product in development at Census longitudinally links the flows of workers across firms, industries, and geographies. One of the firm characteristics available will be the age of the firm (others are geography, industry, and size of firm). Flows of workers across startups will answer numerous questions about the career paths of startup teams, whether startups recruit locally or on the national labor market, whether they workers from nonemployment or poach from established firms in the same industry and how recruiting patterns change over the lifecycle of the firm. This product is currently in development, with a release date of late 2014.[3] In this paper, we will show some sample statistics from the prototype data of flows of workers into and out of startup firms.

These new data can shed light on the causes and consequences of the decline in startups in the United States. Hyatt and Spletzer (2013) have already used LEHD data to show that the shift in the employment away from startups and toward older businesses can account for 5%-10% of the decline in both gross worker flows and job-to-job flows. Documenting the process by which startups accumulate a workforce can inform the question of whether recent changes in the demographics of the United States might deprive startups of the workers they need, or whether the decline of startups might impact certain demographic groups more than others. Tracking what sort of experience makes a person more likely to become an entrepreneur might allow trends in people who have such experience might affect the rate of creation of startups.

Lastly, development is underway to integrate administrative data on self-employed workers into the LEHD data. Previous research has shown that self-employment flows and persistence behave differently than in wage and salary employment, yet no publicly available dataset currently allows researchers and policy-makers to monitor the dynamics and characteristics of this set of potential entrepreneurs on a continual basis. Furthermore, linking the information on self-employment jobs with the UI wage data will shed light on the interaction between two types of economic activity, and even characterize the career paths of entrepreneurs. As described in Garcia-Perez et al. (2012), it is already known that the overlap is substantial, with approximately 50% of the self-employed maintaining a wage and salary job during the same year. The integrated data set will enable the creation of statistics that further describe the interplay between the two universes, and shed light on which self-employed individuals grow employer businesses. Expanding the coverage of LEHD will enable the program to generate data products that richly characterize the transitions into and out of self-employment, as well as the earnings of the business owners, by industry, geography, business age, business size, and owner demographics.

[3] The prototype of this data project is called Job-to-Job Flows and is described in detail in Hyatt and McEntarfer (2012).

Through the development of new and innovative data products, as well as enhancements to existing ones, the U.S. Census Bureau's LEHD program hopes to provide a more comprehensive picture of entrepreneurs and the workforce of startups. In our paper, we will document these recent and ongoing developments and indicate the sort of questions that can be answered with these data. Given these ongoing developments, feedback on these data products is critical to ensure that they are of maximal value to potential data users.

References

Dalton, Sherry, Erik Friesenhahn, James Spletzer, and David Talan. 2011. "Employment Growth by Size Class: Firm and Establishment Data." *Monthly Labor Review* 134(12): 3-12.

Garcia-Perez, Monica, Christopher Goetz, John Haltiwanger, and Kristin Sandusky. 2012. "Don't Quit Your Day Job: Using Wage and Salary Earnings to Support a New Business." Paper presented at the American Economic Association meeting in San Diego, CA in January 2013.

Haltiwanger, John, Henry Hyatt, Erika McEntarfer, and Liliana Sousa. 2012. "Job Creation, Worker Churning, and Wages at Young Businesses." Kauffman Foundation, http://www.kauffman.org/uploadedfiles/bds_report_7.pdf.

Haltiwanger, John, Henry Hyatt, Erika McEntarfer, Kristin Sandusky, and Liliana Sousa. 2013. "Startups are the Gateways to the Labor Market." Unpublished draft, US Census Bureau.

Haltiwanger, John, Javier Miranda, and Ron Jarmin. 2013. "Who Creates Jobs? Small versus Large versus Young." *Review of Economics and Statistics* 95(2): 347-361.

Hyatt, Henry, and Erika McEntarfer. 2012. "Job-to-Job Flows in the Great Recession." "Job-to-Job Flows in the Great Recession." *American Economic Review Papers and Proceedings* 102(3): 580-583.

Hyatt, Henry and James Spletzer. 2013. "The Recent Decline in Employment Dynamics." IZA Discussion Paper No. 7231, http://ftp.iza.org/dp7231.pdf.