Measuring Business Innovation Using a Multi-Dimensional Approach

Lucia Foster, U.S. Census Bureau[1]

Advancing the U.S. Census Bureau's mission "to serve as the nation's leading provider of quality data about its people and economy" requires a robust and agile research and development program working in close collaboration with external experts and Census Bureau programmatic staff. Even straightforward concepts, such as the use of industrial robotics in manufacturing, can require a multi-dimensional measurement approach. While the Census Bureau is known for its surveys, some of our most innovative work combines survey data with administrative data or combines multiple sources of administrative data.

In this brief paper, I discuss the multi-dimensional research and development approach the Center for Economic Studies (CES) at the Census Bureau takes in attempting to better understand business innovation.[2] Since it is not possible to provide details on these many interrelated efforts, I highlight our multi-dimensional approach by giving examples of research using administrative data, survey data, and indirect inference. A more complete view of CES research activities is provided in our annual reports and working paper series.[3]

I. Context

Census is one of thirteen principal statistical agencies in the United States. The missions of these other agencies are often complementary to Census' mission and hence one important activity of CES is outreach to other agencies to partner on topics of mutual interest. When the topic is innovation, we often partner with the National Center for Science and Engineering Statistics (NCSES), but we also partner with other federal agencies, state governments, and other institutions (such as universities). Further, we work with individuals, especially academic experts, to help us improve our measures of the US economy and its people. Many of these researchers conduct work through one of the 30 locations in the Federal Statistical Research Data Center (FSRDC) system.[4] Most of the examples below are based on research conducted with academic experts.

In all of this work, we support U.S.C. Title 13, which allows the use of microdata to provide a benefit to the Census Bureau with conditions to protect the confidentiality of our respondents. Operationally, this pledge of confidentiality may constrain the granularity of publicly available information. For research questions that cannot be answered using published data, researchers can apply to use the data through the FSRDC system.

II. Measuring Business Innovation Using *Administrative* Data

I start by describing two large research and development projects attempting to measure innovation using administrative data: the Business Dynamics Statistics for Patenting Firms (BDS-PF) and the

---

[1] Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results are from existing papers that were reviewed to ensure that no confidential information is disclosed. I thank Emek Basker, John Eltinge, Shawn Klimek, and Nikolas Zolas for helpful comments.

[2] Jarmin (2019) discusses enhancing and improving economic measurement at the Census Bureau.

[3] See https://www.census.gov/programs-surveys/ces/research.html.

[4] See https://www.census.gov/fsrdc.

Innovation Measurement Initiative (IMI). Together they represent the collection and use of administrative data from the federal government, state governments, and universities and demonstrate our collaborations with academic researchers.

The Business Dynamics Statistics (BDS) program provides annual information for the U.S. non-farm economy on firm startups and shutdowns, establishment entry and exit, and job creation and destruction. Core data for the BDS comes from the Census Bureau's business frame which relies heavily on federal administrative data.[5] CES has embarked on a multi-year project to enhance the BDS to include a series of indicators enabling us to provide information on business dynamics by firm characteristics including globalization (exporting, importing, and multinational); human capital (of workers and owners); and innovation (patents, trademarks, R&D expenditures, and other inputs or outcomes of innovative activities). This section focuses on the component of the innovation project identifying firms that patent (BDS-PF).[6]

Multiple research teams have linked patent data to Census business data. An early part of the BDS-PF included a collaboration between the Census Bureau and United States Patent and Trademark Office (USPTO). This team improved upon the existing linkage (previously done through linking assignee information from patent documents to the business register) by incorporating additional inventor information from the same patent documents linked to the Longitudinal Employer-Household Dynamics (LEHD) data. The LEHD data itself relies upon administrative jobs data from state agencies and from federal agencies and data from the Quarterly Census of Employment and Wages provided by states (Abowd et al. (2009)). The researchers triangulate these two independent sources of information (assignees and inventors) to link granted patents to their firm owners, allowing them to substantially improve match rates over earlier studies (Graham et al. (2018)).

The latest research at Census for the BDS-PF focuses on patents related to Artificial Intelligence (AI) and uses natural language processing and machine learning to conduct this research. While the USPTO classifies the technologies embedded within patents according to preexisting classification systems with hundreds of classes and thousands of subclasses, Alderucci et al. (2020) argue that using these and/or keywords will miss much potential AI use since AI is becoming a general purpose technology (GPT). Alderucci et al. (2020) train a machine learning algorithm to identify 52,000 AI-related patents (or up to 140,000 patents using a looser definition) which, they note, is about 3 to 10 times the number of AI patents first identified by Cockburn et al. (2019). This same methodology can potentially be applied across other technology fields.

An entirely different set of metrics comes from the joint Innovation Measurement Initiative (IMI), which links Census data to the Institute for Research on Innovation and Science (IRIS) UMETRICS data from universities on federally sponsored research at the project level. The IRIS data includes project-level financial transactions including payments to internal personnel, payments to outside vendors, and payments to contractors as part of sub-awards. As Lane et al. (2018) note in their overview of the IMI project, the IRIS builds on long running efforts to demonstrate the innovation flowing from federally funded R&D.

[5] Researchers at CES developed the Longitudinal Business Database (LBD) from the business frame (Jarmin and Miranda (2002) and Stinson, White, and Lawrence (2017)) and the BDS is the public product derived from the LBD.
[6] Goldschlag and Perlman (2017) provide an overview of the larger project, Business Dynamics of Innovative Firms.

IRIS currently includes over 30 universities with the goal of partnering with 150 universities (IRIS targets every university with at least $100 million R&D). The data include 392,000 funded awards covering 643,000 research employees and $84 billon in award spending and $61 billion in vendor and subcontract spending.[7] Dissemination of results from this project currently occurs in three ways: research papers, research datasets for qualified users on approved projects, and two quarterly reports (a vendor report and an employee report) at the campus level for participating universities. Additionally, Census and IRIS are developing other publicly available data products.

Researchers have combined the IMI data with Census datasets to examine such subjects as the gender gaps in STEM occupations (Buffington et al. (2016)), outcomes of Ph.D. recipients (Zolas et al. (2015)) and the impact of workers' research experience on new firm outcomes. For the latter, Goldschlag et al. (forthcoming) link the employee data with Census data on startups to look at the link between research experience and young firm outcomes including survival, growth, and innovation. They find that workers research experience is correlated with an "up-or-out" firm dynamic (negatively correlated with survival, but conditional on survival, positively correlated with growth) and with innovative activities (as measured by patent and trademark filings).

III. Measuring Business Innovation Using *Survey* Data

To understand technology adoption and diffusion, we turn to survey data. The Annual Business Survey (ABS) is a relatively new survey (starting with reference year 2018) and represents a partnership between Census and NCSES. This firm-level survey covers all sectors of the non-agricultural economy.[8] The ABS 2018 was mailed to about 850,000 firms (about 560,000 firms responded) and includes sections on innovation (16 questions), technology (three questions) and intellectual property (four questions). My focus is on the three questions in the technology section.

These three questions concern the digital share of business activity (digitization), cloud service purchases, and advanced business technologies for reference year 2017. The digitization question asks firms for the extent certain information types (such as personnel or financial data) are stored in digital format. Similarly, the cloud services purchases question asks firms about which of their IT functions in eight different areas (such as servers and data storage) are stored in the cloud. The third questions asks directly about the testing or use of nine advanced business technologies (for example, machine learning, natural language processing, and robotics).

The survey results suggest that adoption of digitization is widespread, with the use of cloud computing being less so, and adoption of many of the advanced technologies still in their infancy (Zolas et al. (2020)). We find nearly 70% of the firms have adopted some form of digitization (mainly for personnel and financial information), while more than 50% of firms report either no cloud purchases or that they are not necessary. Turning to advanced technologies, we find that 2.2% of respondents are using machine learning and less than one percent are testing its use.

Looking forward, the ABS 2019 has two sections especially relevant for innovation. The "Products and Processes" section has nine questions concerning new or improved goods, services, and business processes. Follow-up questions further distinguish between "new to the business" and "new to the

---

[7] For more information see: https://iris.isr.umich.edu.

[8] See also the Business Research and Development and Innovation Survey and the Annual Survey of Entrepreneurs.

market". The "Technology and Workforce" section has 32 questions about workforce composition and demand and five advanced technologies, including AI, robotics and specialized software. Researchers interested in innovation may find the question concerning factors prohibiting technology adoption and utilization in production especially interesting. At the time of this writing, the responses from the 300,000 firms surveyed have been collected and are being processed.

IV. Applying *Indirect Inference* to Identify Innovative Activities

Given the challenges associated with measuring innovation directly, the last approach relies on indirect inference to identify areas in the economy with innovative activity. Using micro-level data on productivity growth and business entry and exit, Foster et al. (forthcoming) identify patterns in these dynamics that are suggestive of innovative activity. We build on the stages of firm dynamics in response to innovation developed by Gort and Klepper (1982) which we summarize as: innovation leads to a burst of business entry, which is followed by experimentation and adoption, and ultimately a period in which businesses who have successfully responded to the innovation grow while those that have not, shrink and exit.

Foster et al. (forthcoming) apply findings from the literature on the importance of reallocation for aggregate productivity growth to these stages of firm dynamics. Thus, following an innovation, we expect to see business entry, which leads to productivity dispersion as businesses experiment, then rising productivity growth as some businesses becoming more productive and resources reallocate towards successful businesses. Eventually, productivity dispersion compresses as the sector matures and settles down. Their analysis comparing outcomes of high-tech versus non-high-tech industries suggests that these patterns may be useful guides when looking for industries with innovation.

V. Conclusion

The Census microdata referenced in this paper are available for qualified researchers on approved projects use through the FSRDC system. The CES Working Paper Series and Technical Working Paper series include many papers documenting various Census surveys and research data sets. As this short paper has made clear, the Census Bureau leverages its partnership with academic experts in order to continually improve our measures of our nation's people and economy. Understanding innovation is a critical component in this work and perhaps this short paper will inspire more researchers to utilize the FSRDC network to help us better understand business innovation.

References

Abowd, John M., Bryce E. Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L. McKinney, Marc Roemer, and Simon D. Woodcock (2009), "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators." In Producer Dynamics: New Evidence from Micro Data, Vol. 68, Studies in Income and Wealth, eds. Timothy Dunne, J. Bradford Jensen, and Mark J. Roberts. Chicago: University of Chicago Press.

Alderucci, Dean, Lee Branstetter, Eduard Hovy, Andrew Runge, and Nikolas Zolas (2020), "Quantifying the Impact of AI on Productivity and Labor Demand: Evidence from U.S. Census Microdata," mimeo.

Buffington, Catherine, Benjamin Cerf, Christina Jones, and Bruce Weinberg (2016), "STEM Training and Early Career Outcomes of Female and Male Graduate Students: Evidence from UMETRICS Data Linked to the 2010 Census," *American Economic Review: Papers and Proceedings*, 106:333–338.

Cockburn, Iain, Rebecca Henderson, and Scott Stern (2019), "The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis," in, (Eds.), The Economics of Artificial Intelligence: An Agenda, eds. Ajay Agrawal, Joshua Gans, and Avi Goldfarb. Chicago: University of Chicago Press.

Foster, Lucia, Cheryl Grim, John Haltiwanger, and Zoltan Wolf (forthcoming), "Innovation, Productivity Dispersion, and Productivity Growth," in *Measuring and Accounting for Innovation in the 21st Century*, eds. Daniel Sichel, Carol Corrado, Jonathan Haskel, and Javier Miranda, University of Chicago Press.

Goldschlag, Nathan and Elisabeth Perlman (2017), "Business Dynamics of Innovative Firms," Center for Economic Studies Working Papers 17-72, U.S. Census Bureau.

Goldschlag, Nathan, Ron Jarmin, Julia Lane, and Nikolas Zolas (forthcoming), "Research Experience as Human Capital in New Business Outcomes," in *Measuring and Accounting for Innovation in the 21st Century*, eds. Daniel Sichel, Carol Corrado, Jonathan Haskel, and Javier Miranda, University of Chicago Press.

Gort, Michael and Steven Klepper (1982), "Time Paths in the Diffusion of Product Innovations," The Economic Journal, 92(367): 630-653.

Graham, Stuart. J., Cheryl Grim, Tariqul Islam, Alan C. Marco, and Javier Miranda (2018) "Business Dynamics of Innovating Firms: Linking US patents with Administrative Data on Workers and Firms," *Journal of Economics and Management Strategy*, 27(3), 372-402.

Jarmin, Ron S. (2019), "Evolving Measurement for Evolving Economy: Thoughts on 21st Century US Economic Statistics," *Journal of Economic Perspectives,* 33(1), 165-184.

Jarmin, Ron and Javier Miranda (2002), "The Longitudinal Business Database," Center for Economic Studies, Working Paper 02-17, U.S. Census Bureau.

Lane, Julia, Jason-Owen Smith, Joseph Staudt, and Bruce Weinberg (2018), "New Measurement of Innovation" in Center for Economic Studies and Research Data Centers Research Report: 2017, U.S. Census Bureau.

Stinson, Martha, T. Kirk White, and James Lawrence (2017) "Upcoming Improvements to the Longitudinal Business Database and the Business Dynamics Statistics," mimeo.

Zolas, Nikolas, Zachary Kroff, Erik Brynjolfsson, Kristina McElheran, David Beede, Catherine Buffington, Nathan Goldschlag, Lucia Foster, Emin Dinlersoz (2020) "Advanced Technologies Adoption and Use by U.S. Firms: Evidence from the Annual Business Survey", mimeo.

Zolas, Nikolas, Nathan Goldschlag, Ron Jarmin, Paula Stephan, Jason Owen-Smith, Rebecca F. Rosen, Barbara McFadden Allen, Bruce A. Weinberg, and Julia I. Lane (2015), "Wrapping it up in a Person: Examining Employment and Earnings Outcomes for Ph. D. Recipients." *Science* 350(6266): 1367-1371.