

Research & Knowledge



Linked micro-data: A goldmine for economic research and policy making

BY ESTELLE CANTILLON

Would you buy shares in a global firm by looking at its activities in only one country? Of course not! Yet, administrative firm-level data are largely organized across national lines, with different identification systems that make it difficult to link crucial information about their operations across different countries. Fortunately, this is changing to the benefit of research and policy. But more should be done.

At the policy-level, the wake-up call was the onset of the 2008 financial crisis. National regulators struggled to understand the real exposure of their banks because their data structures, based on domestic legal entities, did not capture the complexity of the operations of these banks. Lehman Brothers, for example, consisted of thousands of legal entities, with operations in many countries. This lack of appropriate data complicated the planning for the Lehman Brothers bankruptcy and slowed down the policy response to the crisis.

Since then, national regulators and international institutions such as the Financial

Stability Board and the Bank for International Settlements have worked to improve financial statistics as part of the G20 Data Gaps Initiative. National regulators share their data, and financial interconnections are now better measured and better understood than 10 years ago.

New insights from linked firm-level data



The benefit from linking business data across activities, legal entities and national

boundaries extend beyond the financial sector. For example, one of the most important recent research findings in trade is that intra-firm trade, i.e. international shipments of goods and services between two different units of the same firm, represents a sizable proportion of overall trade transactions (Lanz and Miroudot, 2011). This finding has huge implications for policy because the price used for these transactions are internal transfer prices and not market prices. As such, they are driven by other considerations such as taxation and profit-shifting. Another example is the ability to better measure productivity by linking firm-level employment data with firm-level sales data, but this is only possible if the same firm has the same identifier in different administrative datasets.

There is no question that data quality is improving steadily on this front. Beyond traditional data vendors, such as Bureau Van Dijk, that work on linking publicly available

business data, national statistical systems are revamping the way they collect and organize data. Unique identifiers to store data on the different activities of firms are becoming the norm. There are also several initiatives by national statistical agencies to explore the data needs of linking business data across borders. Routine data linking across borders might become a reality in 10 years.

Linked individual data



Linking administrative data on individuals is also proving to be very valuable. Indeed, traditionally, researchers have relied on survey data to analyze the circumstances and behavior of individuals, but linked individual administrative data have the benefit that they cover the entire population and that they track it over time. We can now better understand questions such as path dependency in education, how careers develop as a function of individual circumstances (health, family status), or the drivers of social mobility. Linked microdata provide boundless opportunities for new research questions. For example, Dahl et al. (2012) used Danish data linking CEOs' family characteristics with pay levels and gender wage differences in their firms. They found that *average* salaries tend to go down when male CEOs father a child but the effect is lower when that child is a girl, and that gender wage differences tend to decrease. In the UK, linking data from the National Pupil Database with higher education records has helped shed new light on an old debate about whether university admissions are socially-biased by showing that prior academic performance was the main determinant of admission decisions (Chowdry et al. 2013).

Nordic leadership

The problem, of course, is that microdata are sensitive. Individuals have the right to have

their privacy protected. Firm-level data may contain commercially-sensitive information. Providing secure access to linked microdata is also resource-intensive. The new EU data protection regulation allows the use of microdata for research purposes but continues to devolve its implementation to Member States. The result is a patchwork of national rules and norms where Nordic countries stand out as world leaders in providing access to microdata for research. Their success is based on a long tradition of collecting data for administrative and statistical purposes and on some of the highest level of protection of personal and business data. Projects are carefully screened to evaluate their societal interest and to ensure the data request is legitimate given the research question. Researchers must be affiliated with a research institution that guarantees the technical protection of the data and takes responsibility for any violation of confidentiality. This creates the necessary trust for public support and a shared understanding that society benefits when researchers can access these data.



Other countries are currently developing their own models of access to microdata, notably Germany and the United Kingdom. In some, specific data owners may take the initiative of facilitating access for research (this is the case of the Belgian National Bank for example). In many countries however, access remains impossible or extremely long and arduous.

Establishing Europe as a data powerhouse for research and policy

ULB was part of the [COEURE Coordination Action](#) funded by the European Commission to map the research frontier and remaining challenges for economic research in Europe.



In their manifesto, COEURE calls for establishing Europe as a data power-house for research by facilitating data access for researchers, improving data design and harmonization, and supporting data infrastructure in Europe. Data are a goldmine for research. Europe, a collection of countries with different economic regulations and laws, offers incomparable opportunities for original research and new policy-relevant insights into the workings of our economies. Making Europe a data powerhouse means providing European-based researchers an edge for generating innovative and impactful research and supporting evidence-based policy-making. All stakeholders have a role to play in making this vision a reality. **Governments** need to promote data access for scientific purposes by lowering non legal barriers and give statistical agencies and central banks a mandate to service researchers. **Statistical agencies and central banks** need to work on data harmonization and linking. **Research institutions** need to establish protocols to guarantee the integrity of the data entrusted to their researchers and build the ethical, legal and technical expertise to support them. Last but not least, **researchers** can build trust and support for greater data access by showcasing the value of their research based on such data.

Note: This article is based on Chapter 13: Development in Data for Economic Research of the book "Economics without Borders" to be published by Cambridge University Press in January 2017 (preprint version available at: <http://www.coeure-book.ceu.edu/>)

References

- Chowdry, H., C. Crawford, L. Dearden, A. Goodman and A. Vignoles 2013. Widening Participation in Higher Education: Analysis using Linked Administrative Data, *Journal of the Royal Statistical Society Series A, Royal Statistical Society*, **176**(2), 431-457.
- Dahl, Michael, Cristian Deszo and David Gaddis Ross 2012. Fatherhood and Managerial Style: How a Male CEO's Children Affect the Wages of His Employees, *Administrative Science Quarterly*, **57**(4), pp. 669-693.
- Lanz, R. and S. Miroudot 2011. Intra-Firm Trade: Patterns, Determinants and Policy Implications, OECD Trade Policy Papers, No. 114.
-



Estelle Cantillon

Estelle Cantillon is FNRS research director at ECARES and Associate Dean for Academic Programmes at SBSEM. She teaches environmental economics at the Masters level. Her research lies at the boundary between market design and industrial organization and typically combines theory and data. She holds a Ph.D. in Economics from Harvard University and has held faculty appointments at Yale University, Harvard Business School and the Harvard Kennedy School of Government.

Contact: Estelle.Cantillon@ulb.ac.be