



## NEW REPORT: The challenges of Big Data: Economists need to acquire new skills to take full advantage of the volume of data sets available

Due to the prevalence of connected digital devices, observational data sets are now much larger and of higher frequency than traditional surveys. A [new IZA World of Labor report](#) shows researchers need fresh training and thinking to learn from data sets of this size.

The rise of Big Data is an exciting time for the ambitious economist, policymaker, or social scientist. Never before has so much data been available to test existing theories and develop new ones. Economists have a natural edge in this endeavor, as they are used to working with complex data. However, this edge is rapidly declining. While in small data sets traditional econometric methods tend to outperform more complex techniques, in large data sets machine learning methods shine. Therefore, according to the economists Matthew Harding, University of California, and Jonathan Hersh, Chapman University, new analytic approaches are needed to make the most of Big Data in economics.

One of the challenges of Big Data is having to manage larger data sets with many, even thousands of, variables. Without a clear prior understanding of the underlying data generating process to direct one's effort, time can be wasted inefficiently searching through options.

Furthermore, in the model fitting process, parameters of machine learning methods often need to be calibrated, or tuned. And to do so, economists need to understand the processes. According to Harding and Hersh, using Big Data or machine learning does not mean that an economist needs to approach an analysis from a radically different viewpoint. Very often, machine learning tools simply enhance the existing econometric methodology.

Harding and Hersh give an example of machine learning techniques in a recent study from 2017. The aim was to construct a structural food demand model and simulate the impact of different product and nutrient taxes in the US. The richness of the existing food transaction data (barcodes) means that food purchases are observed for over 1.1 million distinct food items, an impossible set of data to analyze with traditional economic methods. This is where machine learning can help, by providing a more robust approach using an algorithm to cluster products into distinct groups based on their detailed nutrition profiles (e.g. calories, fat, and sugar). But even this simple algorithm needed fine tuning by the economists undertaking the research.

Harding and Hersh conclude: *"Newer methods from machine learning are expanding the ability to handle Big Data at scale, and researchers risk being cut off from the frontier if these methods are not incorporated into their toolkit. Economists understand how to construct and test causal statements, which makes their skills highly valuable in a data-saturated world. The challenge lies in learning how to implement these methods at scale."*

*Please credit IZA World of Labor should you refer to or cite from the report.*

Please find further research around digitization and big data on the [IZA World of Labor key topics page](#).



BLOOMSBURY PUBLISHING  
LONDON · NEW DELHI · NEW YORK · SYDNEY

**I Z A**  
**World of Labor**  
Evidence-based policy making

Media Contact:

Please contact Anna von Hahn for more information or for author interviews:  
anna@smallspaceagency.com or +44 7852 882 770

Notes for editors:

IZA World of Labor (<http://wol.iza.org>) is a global, freely available online resource that provides policy makers, academics, journalists, and researchers, with clear, concise and evidence-based knowledge on labor economics issues worldwide.

The site offers relevant and succinct information on topics including diversity, migration, minimum wage, youth unemployment, employment protection, development, education, gender balance, labor mobility and flexibility among others.

Established in 1998, the Institute of Labor Economics ([www.iza.org](http://www.iza.org)) is an independent economic research institute focused on the analysis of global labor markets. Based in Bonn, it operates an international network of about 1,500 economists and researchers spanning more than 45 countries.