WORLD BANK GROUP

RESEARCH NEWSLETTER

Machine Learning and Big Data	**July 2018**	A Note from our Research Director

FEATURE STORY

Can State-of-the-Art Machine Learning Tools Give New Life to Poverty Data?

Machine learning tools are making it possible to use a handful of easy-to-collect pieces of household information to predict the poverty status of a household, allowing for poverty estimation based on simpler household survey instruments. Read More »

RESEARCH HIGHLIGHTS

✓ **Obstacles on the Road to Palestinian Economic Growth**

Roy van der Weide, Bob Rijkers, Brian Blankespoor, and Alexei Abrahams, Policy Research Working Paper 8385, March 2018.

The impact of changes in road infrastructure on local GDP in the West Bank proxied by nighttime lights during deployment of road closure obstacles between 2005 and 2012 suggests GDP per capita would have been 4 to 6 percent higher each year in the absence of obstacles.

✓ **The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees**

Paolo Brunori, Paul Hufe, and Daniel Gerszon Mahler, Policy Research Working Paper 8349, February 2018 | Blog.

Machine learning can provide precise and credible measures of the amount of inequality due to factors beyond individual control.

✓ **Using Satellite Imagery to Assess Impacts of Soil and Water Conservation Measures: Evidence from Ethiopia's Tana-Beles Watershed**

Daniel Ayalew Ali, Klaus Deininger, and Daniel Monchuk, Policy Research Working Paper 8321, January 2018.

Findings from satellite imagery can be used to monitor watershed and land use changes, soil degradation, and impact of interventions in real time at a fraction of the cost.

✓ **Mapping the world population one building at a time**

Tobias G. Tiecke, Xianming Liu, Amy Zhang, Andreas Gros, Nan Li, Gregory Yetman, Talip Kilic, Siobhan Murray, Brian Blankespoor, Espen B. Prydz, Hai-Anh H. Dang, Working Paper, Cornell University, December 2017.

A computer vision method is used to create population maps, based principally on the recognition of individual buildings from high-resolution satellite imagery for entire countries, and anchored in and validated with the available census and household survey data.

✔ **Simulation of Synthetic Complex Data: The R Package simPop**

Matthias Templ, Bernhard Meindl, Alexander Kowarik, and Olivier Dupriez, Journal of Statistical Software 79 (Issue 10), August 2017.

simPop is an open source data synthesizer that can combine data from different sources to create "augmented datasets" as inputs to micro-simulation models or other analytical work.

✔ **Man vs. Machine in Predicting Successful Entrepreneurs: Evidence from a Business Plan Competition in Nigeria**

David McKenzie and Dario Sansone, Policy Research Working Paper 8271, December 2017 | Blog.

This research assesses the relative performance of man vs. machine in predicting outcomes for entrants in a business plan competition in Nigeria.

To access the latest Policy Research Working Papers, please click here.

UPCOMING EVENTS

5th Urbanization and Poverty Reduction Research Conference | September 7, 2018

Across the developing world, the growth of cities is outpacing effective policy. Low density land use results in rapidly expanding cities, raising the costs of infrastructure and service provision. At the same time, limited investments in transport infrastructure limit connectivity and the opportunities that make cities engines for growth. Effective policy to address these challenges requires an understanding of the spatial organization of cities. More »

January 15 – 16, 2019 | Globalization: Contents and Discontents (Kuala Lumpur)

For at least twenty years, scholars have debated the pros and cons of globalization. The goal of this conference is to revisit these and other issues to understand the processes underlying globalization and how they may, or may not, shape a better world.

To see more events, please click here.

ANNOUNCEMENTS

The **DIME** wiki is a one-stop shop for resources on all phases of impact evaluation design, fieldwork, data, and analysis.

Wiki | **Blog**

BLOGS

Is Predicted Data a Viable Alternative to Real Data?

The primary motivation for predicting data in economics, health sciences, and other disciplines has been to deal with various forms of missing data problems. However, one could also make a case for adopting prediction methods to obtain more cost-efficient estimates of welfare indicators when it is expensive to observe the outcome of interest. For example, consider the estimation of poverty and malnutrition rates.

Read more »

To read more of our blogs, see: Let's Talk Development | All About Finance | Development Impact