

Towards a Vision for a New Data Infrastructure for Federal Statistics and Social and Economic Research in the 21st Century

Workshop 2: The Implications of Using Multiple Data Sources for Major Survey Programs

May 16 and 18, 2022 Agenda

Monday, May 16, 2022

11:00 – 11:15am EDT

Introduction and Welcome

- **Sharon Lohr**, Consensus Panel Chair, *Arizona State University*
- **Cheryl Eavey**, *National Science Foundation*

11:15 – 1:15 EDT

Session 1: Opportunities for Using Multiple Data Sources to Enhance Major Survey Programs

Moderator: Jean-François Beaumont, *Statistics Canada*

11:20-11:50 Keynote - Census Bureau Modernization: A New Vision for an Enterprise Approach to Statistical Data: **Robert Santos**, *Director, US Census Bureau*

11:50-12:20 Keynote - Combining Multiple Data Sources to Produce Inclusive Official Statistics: **Anil Arora**, *Chief Statistician of Canada*

12:20-12:35 Discussant: **Joseph Salvo**, *University of Virginia*

12:35-12:50 Discussant: **Haoyi Chen**, *United Nations*

12:50-1:15 Panel/Audience Discussion

1:15 – 1:35 EDT

BREAK

Monday, May 16, 2022

1:35 – 3:05 EDT

Session 2: Measuring Crime in the 21st Century: A Panel Discussion

Moderator: Elizabeth Stuart, *Johns Hopkins University*

1:35-2:35

Janet Lauritsen, *University of Missouri-St. Louis*

Ramiro Martinez, Jr., *Northeastern University*

Erica Smith, *US Bureau of Justice Statistics*

Derek Veitenheimer, *State of Wisconsin*

2:35-3:05

Panel/Audience Discussion

3:05 – 3:20 EDT

BREAK

3:20 – 4:50 EDT

Session 3: Improving Agriculture Statistics with New Data Sources

Moderator: Shaowen Wang, *University of Illinois at Urbana-Champaign*

3:25-3:45

The Crops County Estimates Program: Developing Official Statistics Based on Available Data: **Linda Young, *US National Agricultural Statistics Service***

3:45-4:05

The Evolution of the Use of Satellite and Administrative Data in Estimating Mid-season Field Crop Yields at Statistics Canada: **Herbert Nkwimi-Tchahou, *Statistics Canada***

4:05-4:25

Layers and Beyond, Modern Use of Connected Spatial and Non Spatial Datasets to Unlock Insights in R&D: **Martin Mendez-Costabel, *Bayer Crop Science***

4:25-4:35

Discussant: **Michael Goodchild, *University of California, Santa Barbara***

4:35-4:50

Panel/Audience Discussion

4:50 – 5:00 EDT

Wrap-up Day 1 (Sharon Lohr)

5:00pm EDT

Adjournment

Wednesday, May 18, 2022

- | | |
|---------------------|---|
| 11:00 – 11:05am EDT | Introduction and Welcome
Sharon Lohr , <i>Arizona State University</i> |
| 11:05 – 12:40 EDT | Session 4: Data Linkage for Income and Health Statistics
Moderator: Judith A. Seltzer , <i>University of California, Los Angeles</i> |
| 11:10-11:30 | The National Experimental Wellbeing Statistics Project:
Jonathan Rothbaum , <i>US Census Bureau</i> |
| 11:30-11:50 | Realizing the Power of Health Data through Linkages:
Lisa Mirel , <i>US National Center for Health Statistics</i> |
| 11:50-12:10 | Administrative Data Linkages in the Health and Retirement Study: Social Security and Medicare/Medicaid Data:
Jessica Faul & Helen Levy , <i>University of Michigan</i> |
| 12:10-12:20 | Discussant:
Hilary Hoynes , <i>University of California, Berkeley</i> |
| 12:20-12:40 | Panel/Audience Discussion |
| 12:40 – 1:00 EDT | BREAK |
| 1:00 – 3:00 EDT | Session 5: Issues in Data Equity
Moderator: David Mancuso , <i>State of Washington</i> |
| 1:05-1:25 | From Data to Decision-making: Using Data and Engagement to Advance Racial Equity: Steven Brown , <i>Urban Institute</i> |
| 1:25-1:45 | Administrative Data and Statistics for Small Race Groups and Other Populations:
Randall Akee , <i>University of California, Los Angeles</i> |
| 1:45-2:05 | Diversity in Data: How Data Collection Decisions Help and Harm the Outcome: Frauke Kreuter , <i>Ludwig-Maximilians-University of Munich and University of Maryland</i> |
| 2:05-2:15 | Moderator/Discussant questions and reflections:
Margaret Levenstein , <i>University of Michigan</i> |
| 2:15-2:45 | “Fireside Chat”:
Clarence Wardell , <i>Chief Data and Equitable Delivery Officer, Executive Office of the President</i>
Kimberlyn Leary , <i>Harvard University and Urban Institute</i> |
| 2:45-3:00 | Panel/Audience Discussion |
| 3:00 – 3:30 | Discussion and Wrap-up
Moderator: Sharon Lohr , <i>Arizona State University</i>

Robert M. Groves , <i>Chair of Committee on National Statistics, Georgetown University</i> |
| 3:30pm EDT | Adjournment |

**PANEL ON THE IMPLICATIONS OF USING MULTIPLE DATA
SOURCES FOR MAJOR SURVEY PROGRAMS**

Sharon Lohr (Chair), Professor Emerita of Statistics, Arizona State University

Jean-François Beaumont, Senior Statistical Advisor, Statistics Canada

Lawrence D. Bobo, Dean of Social Science and W. E. B. Du Bois Professor of the Social Sciences, Harvard University

Mick P. Couper, Research Professor, Survey Research Center; and Faculty Associate, Populations Study Center, Institute for Social Research, University of Michigan

Hilary Hoynes, Professor of Public Policy and Economics; and Haas Distinguished Chair in Economic Disparities, University of California, Berkeley

Kimberlyn Leary, Associate Professor, Department of Health Policy and Management, Harvard University

David Mancuso, Director of the Research and Data Analysis Division, Washington State Department of Social and Health Services

Judith A. Seltzer, Research Professor of Sociology; and Faculty Fellow California Center for Population Research, University of California, Los Angeles

Elizabeth A. Stuart, Associate Dean for Education and Professor, Department of Mental Health, Department of Biostatistics, Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health

Shaowen Wang, Professor and Head, Department of Geography and Geographic Information Science; Richard and Margaret Romano Professional Scholar, University of Illinois at Urbana-Champaign