

Matthew Sobek

Curriculum Vitae

IPUMS Center for Data Integration
50 Willey Hall, University of Minnesota, Minneapolis, 55455
sobek@umn.edu, (612) 624-5818

Education

- 1997 Ph.D. Department of History, University of Minnesota
1986 B.A. Department of History, University of Louisiana, Lafayette

Professional Positions

- 2001-present Research Scientist, Director of Data Integration, Institute for Social Research and Data Innovation, University of Minnesota
1998-2001 Managing Editor, *Historical Statistics of the United States, Millennial Edition*, Center for Social and Economic Policy, University of California, Riverside
1998 Lecturer, Humphrey Institute of Public Affairs, University of Minnesota
1997-1998 Post-doctoral Research Associate, Department of History, University of Minnesota
1994 Instructor, Department of History, University of Minnesota

External Grants and Contracts

- Microdata for Research on Aging in the Global South. Co-Principal Investigator with Lara Cleveland and Steven Ruggles. National Institutes of Health: NIA R01-AG062601. \$3.1 million. 2019-2023.
- Microdata for Analysis of Early Life Conditions. Co-Principal Investigator with Catherine Fitch and Steven Ruggles. National Institutes of Health: NIA R01-AG041831. \$3.0 million. 2019-2023.
- Microdata for Population Dynamics and Health Research. Co-Principal Investigator with Sarah Flood and Steven Ruggles. National Institutes of Health: NICHD R01-HD043392. \$3.2 million. 2018-2022.
- Integrated Eurasian Microdata Project. Co-Principal Investigator with Lara Cleveland. National Institutes of Health: NICHD R01-HD047283. \$3.0 million. 2016-2021.
- Big Microdata Expansion Project. Co-Principal Investigator with Catherine Fitch and Steven Ruggles. National Institutes of Health: NICHD R01-HD083829. \$2.9 million. 2015-2020.
- Big Data for Population Research. Co-Principal Investigator with Catherine Fitch and Steven Ruggles. National Institutes of Health: NICHD R01-HD078322. \$3.0 million. 2013-2018.
- Integrated Samples of Latin American Censuses. Co-Principal Investigator with Lara Cleveland. National Institutes of Health: NICHD R01-HD044154. \$3.1 million. 2013-2019.
- Survey Data Harmonization. Principal Investigator, sub-contract with NORC; Felicia LeClere PI. Food and Drug Administration, Center for Tobacco Products. \$1.5 million. 2015-2017.
- Causal Factors in Child Growth: Evidence from Aggregated Survey Data. Principal Investigator. Bill & Melinda Gates Foundation. \$450 thousand. 2015-2017.
- Data Integration and Harmonization. SESTAT Science and Engineering Education Data. Principal Investigator. National Institutes of Health via NETE. \$460 thousand. 2015-2016.

American Time Use Survey: Data Access System. Principal Investigator, subaward with University of Maryland; Katharine Abraham PI. National Institutes of Health: NICHD R01-HD053654. \$1.2 million. 2006-2011.

Tanzania Census Microdata Sample Project. Principal Investigator. University of Tromso, Norway. \$30 thousand. 2006-2007.

External Grants: Co-Investigator

IPUMS-PMA: Integrated Public Use Microdata Series for PMA2020. Principal Investigator, Elizabeth Boyle. Bill & Melinda Gates Foundation. 2018-2020.

Integrated Demographic and Health Surveys. Principal Investigators, Elizabeth Boyle and Miriam King. National Institutes of Health: NICHD R01-HD069471. 2012-2022.

IPUMS-Redesign. Principal Investigator, Steven Ruggles. National Institutes of Health: NICHD R01-HD43392. 2002-2007, 2013-2018.

Microdata for Analysis of Early Life Conditions, Health, and Population. Principal Investigator, Steven Ruggles. National Institutes of Health: NIA R01-AG041831. 2012-2017.

Baseline Socioeconomic Microdata for Population and Health Research. Principal Investigator, Steven Ruggles. National Institutes of Health: NICHD R01-HD073967. 2012-2017.

International Time Use Data Access System. Principal Investigator, Sandra Hofferth. National Institutes of Health: NICHD R01-HD053654. 2012-2017.

Integrated Samples of Eurasian Censuses. Principal Investigator, Robert McCaa. National Institutes of Health: NICHD. 2004-2015.

International Integrated Microdata Series. Principal Investigator, Steven Ruggles. National Science Foundation SES-0433654. 2009-2014.

Integrated Samples of Latin American Censuses. Principal Investigator, Robert McCaa. National Institutes of Health: NICHD R01-HD044154. 2003-2013.

International Integrated Microdata Access System. Principal Investigator, Steven Ruggles. National Science Foundation SBR-9908380. 1999-2004.

Publications

Matthew Sobek. Forthcoming 2020. "The IPUMS Approach to Harmonizing the World's Population Census Data." In Akinori Okada, Sadaaki Miyamoto, Fumitake Sakaori, Yoshiro Yamamoto and Maurizio Vichi, eds., *Advanced Research in Classification and Data Science*. Springer.

Matthew Sobek. 2016. "Data Prospects: IPUMS-International." In Michael White, ed., *International Handbook of Migration and Population Distribution*. New York: Springer, 157-174.

Steven Ruggles, Robert McCaa, Matthew Sobek, and Lara Cleveland. 2015. "The IPUMS Collaboration: Integrating and Disseminating the World's Population Microdata." *Journal of Demographic Economics* 81 (June): 203-216.

Matthew Sobek. 2012. "Making Social Class Work." *Historical Methods* 45 (Fall): 187-192.

Lara Cleveland, Robert McCaa, Steven Ruggles, and Matthew Sobek. 2012. "When Excessive Perturbation Goes Wrong and Why IPUMS-International Relies Instead on Sampling, Suppression, Swapping and Other Minimally Harmful Methods to Protect Privacy of Census Microdata." In J. Domingo-Ferrer and I. Tinnirello (eds.), *Privacy in Statistical Databases 2012*, LNCS. Heidelberg: Springer, 179-187.

- Matthew Sobek, Lara Cleveland, Sarah Flood, Patricia Kelly Hall, Miriam L. King, Steven Ruggles, and Matthew Schroeder. 2011. "Big Data: Large-Scale Historical Infrastructure from the Minnesota Population Center." *Historical Methods* 44 (April-June): 61-68.
- Steven Ruggles, Evan Roberts, Sula Sarkar, and Matthew Sobek. 2011. "The North Atlantic Population Project: Progress and Prospects." *Historical Methods* 44 (January-March): 1-6.
- Robert McCaa, Steven Ruggles, and Matthew Sobek. 2010. "IPUMS-International Statistical Disclosure Controls: 159 Census Microdata Samples in Dissemination, 100+ in Preparation." In J. Domingo-Ferrer and E. Magkos, eds., *Privacy in Statistical Databases 2010*, LNCS 6344. Heidelberg: Springer, 74-84.
- Robert McCaa, Albert Esteve, Steven Ruggles, and Matthew Sobek. 2006. "Using Integrated Census Microdata for Evidence-Based Policy Making: The IPUMS-International Global Initiative." *African Statistical Journal* 2: 83-100.
- Susan Carter et al., editors. 2006. *Historical Statistics of the United States: Millennial Edition*. [Matthew Sobek, Managing Editor]. New York: Cambridge University Press.
- Matthew Sobek. 2006. "Occupations." In Susan Carter et al., eds., *Historical Statistics of the United States: Millennial Edition*. New York: Cambridge University Press.
- Robert McCaa, Steven Ruggles, and Matthew Sobek. 2005. "Integrated Census Microdata Samples for Scientific Research and Teaching." *Estrala* 8 (August): 25-31.
- Robert McCaa, Albert Esteve, Steven Ruggles, and Matthew Sobek. 2005. "La Integración de los Microdatos Censales de América Latina: El Proyecto IPUMS." *Estudios Demográficos y Urbanos* 20 (Spring): 37-70.
- Albert Esteve and Matthew Sobek. 2003. "Challenges and Methods of International Census Harmonization." *Historical Methods* 36 (Spring): 66-79.
- Steven Ruggles, Miriam King, Deborah Levison, Robert McCaa and Matthew Sobek. 2003. "IPUMS-International." *Historical Methods* 36 (Spring): 60-65.
- Steven Ruggles, Matthew Sobek, Miriam L. King, Carolyn Liebler, and Catherine Fitch. 2003. "IPUMS Redesign." *Historical Methods* 36 (Winter): 9-20.
- Matthew Sobek. 2001. "New Statistics on the U.S. Labor Force, 1850-1990." *Historical Methods* 34 (Spring): 71-87.
- Steven Ruggles, Catherine Fitch, Patricia Kelly Hall, and Matthew Sobek. 2000. "IPUMS-USA: Integrated Public Use Microdata for the United States." In Patricia Kelly Hall et al., eds., *Handbook of International Historical Microdata for Population Research*. Minneapolis: Minnesota Population Center, 259-284.
- Matthew Sobek, and Steven Ruggles. 1999. "The IPUMS Project: An Update." *Historical Methods* 32 (Summer): 102-110.
- Todd Gardner, Matthew Sobek, and Steven Ruggles. 1999. "IPUMS Data Extraction System." *Historical Methods* 32 (Summer): 119-124.
- Matthew Sobek. 1997. *A Century of Work: Gender, Labor Force Participation, and Occupational Attainment in the United States, 1880-1990*. PhD Dissertation. University of Minnesota.
- Steven Ruggles, Matthew Sobek, and Todd Gardner. 1997. "Distributing Large Historical Census Samples on the Internet." *History and Computing* 8 (Summer): 145-159.

- Matthew Sobek. 1996. "Work, Status, and Income: Men in the American Occupational Structure since the Late Nineteenth Century." *Social Science History* 20 (Summer): 169-207.
- Steven Ruggles, Todd Gardner, and Matthew Sobek. 1996. "Disseminating Historical Census Data on the World Wide Web." *IASSIST Quarterly* 20 (Fall): 4-18.
- Matthew Sobek. 1995. "The Comparability of Occupations and the Generation of Income Scores." *Historical Methods* 28 (Spring): 47-51.
- Matthew Sobek and Lisa Dillon. 1995. "Interpreting Work: Occupational Classification in the Historical Public Use Microdata Samples." *Historical Methods* 28 (Spring): 70-73.
- Steven Ruggles, J. David Hacker, and Matthew Sobek. 1995. "Order Out of Chaos: General Design of the Integrated Public Use Microdata Series." *Historical Methods* 28 (Spring): 33-39.
- Matthew Sobek. 1991. "Class Analysis and the U.S. Census Public Use Samples." *Historical Methods* 24 (Fall): 171-181.

Working Papers

- Matthew Sobek and Lara Cleveland. 2017. "IPUMS Approach to Harmonizing International Census and Survey Data." *United Nations Economic Commission for Europe, Conference of European Statisticians, Working Paper 31*.
- Matthew Sobek and Sheela Kennedy. 2009. "The Development of Family Interrelationship Variables for International Census Data." *Minnesota Population Center Working Paper Series, 2009-02*.
- Matthew Sobek, Monty Hindman, and Steven Ruggles. 2007. "Using Cyber-Resources to Build Databases for Social Science Research." *Minnesota Population Center Working Paper Series, 2007-01*.

Published Datasets

- Minnesota Population Center. *Integrated Public Use Microdata Series, International: Version 7.2*. [dataset]. Minneapolis, MN: IPUMS, 2019. <https://doi.org/10.18128/D020.V7.2>.
- Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. *IPUMS USA: Version 9.0*. [dataset]. Minneapolis, MN: IPUMS, 2019. <https://doi.org/10.18128/D010.V9.0>.
- Elizabeth Heger Boyle, Miriam King, and Matthew Sobek. *IPUMS-Demographic and Health Surveys: Version 7* [dataset]. Minnesota Population Center and ICF International, 2019. <https://doi.org/10.18128/D080.V7>
- Sandra Hofferth, Sarah Flood, and Matthew Sobek. *American Time Use Survey Data Extract Builder: Version 2.7*. [dataset]. College Park, MD: University of Maryland and Minneapolis, MN: IPUMS, 2018. <https://doi.org/10.18128/D060.V2.7>

Conference Presentations

- "Analyzing Policy Challenges with Real World Data Using IPUMS." With Sula Sarkar. *Asian Population Association*, Shanghai, July 2018.
- "Creating a National Longitudinal Research Infrastructure." With Steven Ruggles, J. Trent Alexander, John Warren, and Catherine Fitch. *International Union for the Scientific Study of Population*, Cape Town, November 2017.
- "The IPUMS Approach to Harmonizing the World's Population Data." *Conference of the International Federation of Classification Societies*, Tokyo, August 2017.

- “Terra Populus: Population and Environment Data.” *United Nations World Data Forum*, Cape Town, January 2017.
- “Preservation and Data Access Efforts by the IPUMS Census Project and its NSO Partners.” *Union for African Population Studies*, Pretoria, South Africa, November 2015.
- “Harmonized Census Geography and Spatio-Temporal Analysis: Gender Equality and Empowerment of Women in Africa.” With Sula Sukar, Lara Cleveland, and Majory K. Silisyene. *Population Association of America*, San Diego, April 2015.
- “Terra Populus: Integrated Data on Population and Environment.” With Tracy Kugler and David Van Riper. *Population Association of America*, Boston, May 2014.
- “Big Microdata from the U.S. Census.” With Steven Ruggles and Catherine Fitch. *Population Association of America*, Boston, May 2014.
- “Terra Populus: Integrated Data on Population and Environment.” With Steven Ruggles, Catherine Fitch, Jonathan Foley, Tracy Kugler, Steven Manson, Peter Clark, and David Van Riper. *International Union for the Scientific Study of Population*, Busan, Korea, August 2013.
- “The IPUMS-International Census Microdata Project.” With Patricia Kelly Hall and Lara Cleveland. [Poster presentation]. *International Statistical Institute Meetings*, Hong Kong, China, August 2013.
- “Harnessing the use of microdata for migration research: the IPUMS-International database” With Robert McCaa. *Eleventh Coordination Meeting on International Migration*, United Nations Population Division, New York, February 2013.
- “Terra Populus: A Global Population/Environment Data Network.” With Steven Ruggles, Catherine Fitch, Jonathan Foley, and Steven Manson. *Population Association of America Meetings*, San Francisco, May 2012.
- “Identifying Family Relationships for the World's Largest Collection of Census Microdata.” With Sheila Kennedy. [Poster presentation]. *Population Association of America Meetings*, Dallas, April 2010.
- “Harmonizing the World's Census Microdata: Methods and Challenges of the IPUMS Project.” *International Statistical Institute Meetings*, Durban, South Africa, August 2009.
- “International Census Data for Comparative Research.” *European Population Conference*, Barcelona, Spain, July 2008.
- “Challenges of Census Data Harmonization.” *Joint Statistics Meetings*, Minneapolis, October August 2005.
- “The IPUMS-International Project: Challenges and Methods of International Census Data Integration.” With Albert Esteve and Robert McCaa. *Social Science History Association*, St. Louis, October 2002.
- “Disseminating Anonymized, Integrated Census Microdata via the Internet: the IPUMS-International Project.” With Robert McCaa and Steven Ruggles. *ANCSDAAP Population Census Conference*, Ulaanbaatar, Mongolia, June 2002.
- “Family Labor Patterns and Economic Status in the United States, 1880-1990.” *Social Science History Association*, Fort Worth, November 1999.
- “High Education, High Technology, and High Wages: An Exploration into the Relationship Between Research Universities and Economic Prosperity and Dynamism.” With Susan Carter and Richard

- Sutch. *R&D Investment and Economic Growth in the 20th Century*, All-UC Group in Economic History, Berkeley, March 1999.
- “Work, Gender, and the U.S. Census: The 1910 ‘Overcount’ of Female Labor.” With Diana Magnuson. *Social Science History Association*, Washington D.C., October 1997. Another version presented at the *Upper Midwest Conference on Demographics for Policy Makers*, St. Paul, April 1998.
- “Opportunities for Developing New Time Series Using the Historical Census Samples.” *Historical Statistics on Labor, Human Capital, and Labor Markets*. National Science Foundation Conference, Palo Alto, June 1997.
- “Microdata and Economic History.” Roundtable: New Directions in Economic History, *Social Science History Association*, Chicago, November 1995.
- “Electronic Dissemination of Historical Census Data.” *Population Association of America*, New Orleans, May 1996. Other versions of this paper presented at *Social Science History Association*, New Orleans, October 1996; *International Association for Social Sciences Information Service and Technology*, Minneapolis, April 1996; *Association of Public Data Users*, Washington D.C., October 1997.
- “Occupational Classification and Measurement: Issues for Historical Analysis Using the Integrated Public Use Microdata Series.” *All-UC Conference on the Public Use Microdata Samples*, Riverside, March 1995.
- “Work, Status, and Income since the Late Nineteenth Century.” *Social Science History Association*, Atlanta, October 1994.
- “America at 1880: Research from the Public Use Sample.” Roundtable Discussant. *Social Science History Association*, Baltimore, November 1992.
- “Occupation, Class, and Socioeconomic Status in Historical Census Data.” *Social Science History Association*, New Orleans, November 1991.
- “Class Analysis and the U.S. Census Public Use Samples.” *Social Science History Association*, Minneapolis, October 1990.

Seminars and Workshop Presentations

- “The Data Revolution in Action: National and International Experiences with Microdata Dissemination and Public Use.” Roundtable presentation. Side meeting of 49th Session of the UN Commission on Population and Development, New York City, April 2016.
- “Workshop on Demographic Data and Techniques, with Resources and Methods from IPUMS and REVES.” Mekong Region Development Research Group, Ho Chi Minh City, Vietnam, May 2015.
- “The Minnesota Data Harmonization Projects.” Bill and Melinda Gates Foundation, Seattle, May 2014.
- “Data Projects at the Minnesota Population Center: Resources for Comparative Population and Health Research.” Center for Studies in Demography and Ecology, University of Washington, Seattle, May 2014.
- “Data Harmonization.” FDA Tobacco Products Scientific Advisory Committee Meeting, Washington DC, April 2014.
- “Use of the IPUMS-International Data Series.” Workshop. International Statistical Institute Meetings, Durban, South Africa, August 2009.

- “The Harmonization of International Census Microdata for Demographic Research: The IPUMS Project.” With Albert Esteve. INED Seminar Series, Paris, March 2009.
- “Uses and Limitations of the IPUMS-International Database.” Seminar. UN Population Council, Cairo, October 2008.
- “Design and Use of the IPUMS-International Census Data Series.” Social Sciences Seminar, Cairo University, October 2008.
- “Workshop: Design and Use of the IPUMS International Census Database.” Fifth African Population Conference, Arusha, Tanzania, December 2007.
- “The Integrated Public Use Microdata Series.” African Census Workshop, Dar es Salaam, Tanzania, January 2007.
- “Using the Integrated Public Use Microdata Series in Research.” American Sociological Association, August 2006 [earlier versions 2003, 2004, 2005].
- “Data Processing in the IPUMS-International Data Series.” European Census Harmonization Conference, Paris, France, June 2006.
- “IPUMS-International Integration Methods and Data Enhancements.” Workshop: Integrating European Census Microdata, Barcelona, Spain, July 2005.
- “Teaching with the Integrated Public Use Microdata Series.” Midwest Sociological Society Meetings, Minneapolis, April 2005.
- “IPUMS-Latin America: Work Process and Calendar.” IPUMS-Latin America: Harmonization and Dissemination of Census Microdata, Cartagena, Colombia, January 2004.
- “IPUMS-Latin America: Dataset Processing.” IPUMS-Latin America: Harmonization and Dissemination of Census Microdata, Cartagena, Colombia, January 2004.
- “Using the Integrated Public Use Microdata Series in Research.” ICPSR Summer Workshop, July 2002.
- “Historical Patterns of U.S. Female Labor Force Participation.” Research seminar, Department of Economics, University of Minnesota, March 2002.
- “Historical Measures of Female Labor Force Participation.” University Lecture Series. University of Louisiana at Lafayette, February 2002.
- “Using the Integrated Public Use Microdata Series.” National Computer-Assisted Reporting Conference, Lexington, KY, September 2000.
- “Measures of U.S. Female Labor Force Participation, 1880-1990.” Seminar. Department of Economics, University of California, Riverside, May 2000.
- “Taking the Measure of Women's Work: Female Labor Force Participation in the U.S. Census, 1880-1990,” Population Research Center, University of Texas at Austin, February 1999.
- “Using U.S. Census Data to Study the Elderly,” Population Studies Center, University of Michigan, July 1996.

Professional Service

Executive Committee, *Whole Village Project*, 2009-2012

Editorial board, *Historical Methods* journal, 2006-2008

Manuscript reviewer for various journals

Proposal reviewer, National Science Foundation, Netherlands Organisation for Scientific Research

Honors and Awards

Schoonover Fellowship, University of Minnesota, 1987-1988

University Fellowship, University of Louisiana at Lafayette, 1986-1987

Walter Craddock Memorial Scholarship, University of Louisiana at Lafayette, Fall 1985

University Academic Scholarship, University of Louisiana at Lafayette, 1981-1982

Membership in Professional Organizations

Population Association of America

International Union for the Scientific Study of Population

Union for African Population Studies

Asian Population Association

Social Science History Association