

CPRDC

CENTRAL PLAINS RESEARCH DATA CENTER

Newsletter, June 2020

COVID-19 RESEARCH OPPORTUNITIES USING RDC DATA

The Central Plains Research Data Center (CPRDC) provides access to numerous restricted federal data sets that may be helpful to researchers interested in pursuing COVID-19 related topics. Here is a brief outline of some of the possibilities.

RDC Resources for Covid-19 Research

The Federal Statistical Research Data Center Network provides qualifying researchers with access to a remarkable range of restricted data sets from across many federal agencies. These data sets provide unique opportunities to examine most topics in the social sciences, economics, and public health, and the impacts of Covid-19 are no exception.

At present, the data does not yet exist to observe the direct effects of the 2020 Pandemic. (The Federal Government is continuing most of its ongoing surveys despite the present crisis, but these results take time to compile and process.) But the existing data can be used in two broad ways: (i) combining outside (state or private) data sets, about infections, transmission, et cetera, with Census data to learn more about details about the sample and (ii) examining past micro-data held by the Federal Government to examine medical, social, about infections, transmission, et cetera, with Census data to learn more about details about the sample and (ii) examining past about infections, transmission, et cetera, with Census data to learn about infections, transmission, et cetera, with examining past micro-data held by the Federal Government to examine



medical, social, and economic questions and models motivated by the Pandemic and its policy response.

Census is often happy to gain knowledge from new data resources from state or local government or even private sources. If a researcher had access to some Coronavirus-related data they would like to gain deeper knowledge of, it is possible for Census to make certain data available for linking. For one narrow example, suppose a researcher had access to hospital-sourced data about Covid-19 infection outcomes, but only knew individual patient's identities and recent medical procedures



and status. In the RDC this could be matched with several individual-level data-sets such as: (i) the American Community Survey (ACS), which is an enormous data-set sampling more than 1% of Americans each year, to understand connections between an individual's employment status or occupation and infection outcomes, (ii) the Survey of Income and Program Participation (SIPP), a smaller survey but in panel over a longer time span, to understand connections between long-term employment status or public assistance programs and the demographic disparities of hospitalized infections, or (iii) the Medical Expenditure Panel Survey (MEPS) to understand the interplay between an individual's health insurance and the health care measures used to treat their case of Covid-19. These are merely simple hypothetical examples.

But the SARS-CoV-2 has also raised a host of questions that can be answered by pre-pandemic data. For instance, a researcher could use detailed employee-firm data from the Longitudinal Employer-Household Dynamics (LEHD) to determine if workers in industries we now know are more likely to be harmed by the Pandemic lockdown are more or less likely to frequently change jobs than average (arguably reducing the long-term effect of the lockdown), have more dependents than average (increasing the poverty implications of the lockdown), and the interplay between those two characteristics. Or they could use restricted data in the National Health and Nutrition Examination Survey (NHANES) to examine rarely-considered connections between comorbidities and detailed geography to ask questions like "is the population of people with pre-existing Covid-19 risk-factors more likely, based on geography, to take the subway (a speculated contagion risk) than those without risk-factors?" Or to use detailed manufacturing data from the Annual Survey of Manufactures (ASM) to produce models to predict how shifts from contagion-prone manufacturing methods and industries might change employment and productivity in the future.

These are just a few small examples. (And do not interpret this short write up as a suggestion that any specific hypothetical proposal it outlines would be assured approval for implementation in the RDC.) The opportunities are tremendous, and Census is eager to provide accelerated evaluation of proposals connected to Covid-19. If you have an interest, contact our Census Administrator Seth Kingery.

Contact the CPRDC

John Anderson, Executive Director
310 Whittier Research Center
janderson4@unl.edu
(402) 472-1190



Seth Kingery, Census Administrator
Central Plains Research Data Center
seth.j.kingery@census.gov
(402) 472-0362



The Central Plains Research Data Center (CPRDC) is a special research facility at the University of Nebraska-Lincoln that is operated in partnership with the U.S. Census Bureau to provide researchers with approved projects access to restricted-use federal data. Data are available from a number of federal agencies, including the Census Bureau, the National Center for Health Statistics, the Internal Revenue Service, the Bureau of Labor Statistics, and other agencies. Restricted access data sets are typically more detailed and provide finer geographic coverage than public use data sets. A Census administrator is on site to assist researchers in the development of proposals, to facilitate security clearances, and to assist in implementation of approved research projects. The purpose of the CPRDC is to enhance academic research, research infrastructure, and the recruitment of research oriented faculty and graduate students. The CPRDC is supported by a consortium of university entities including the University of Nebraska-Lincoln, the University of Nebraska Medical Center, Iowa State University, the University of Iowa, and the University of South Dakota.

