Free Registration: Register in advance at this Zoom link.

Session Title: The Use of Data Analytics and Data Science in Policy Analysis

Session Description:
This seminar features four innovative papers that use data analytics and data science for policy evaluation. The first paper discusses a new framework for using DHS administrative data to address basic scientific and policy questions on immigration. The second paper develops five indices to determine which countries should be using taxation to raise funds for the climate bonds issuance and which countries should be granted transfer payments via climate bond premiums. The third paper uses a data science approach to examine how levels of air pollution differ across demographical communities, whether the COVID-19 lockdown differentially affected communities, and offers recommendations to promote environmental justice. The final paper uses the deep learning algorithm Long Short-Term Memory (LSTM) to forecast the macroeconomic impacts of the COVID-19 pandemic on the G-20 countries.

Session Chair: Steve Payson, University of Maryland

Session Organizer: Areerat Kichkha, AIRLEAP

Paper #1

Paper Title: Using DHS Administrative Data to Address Basic Scientific and Policy Questions in Migration Research

Presenter: Guillermina Jasso, New York University

Additional Authors: Mark R. Rosenzweig, Yale University

Abstract:
This paper develops a framework for using DHS administrative data to address basic scientific and policy questions on immigration. The linchpin data set is the time-tested Immigration File, containing basic information for all persons who were granted lawful permanent residence (LPR) in a fiscal year – e.g., information on origin country (countries of birth, nationality, last residence, and chargeability), LPR visa type, previous temporary visa type, and occupation. This data set is well known, having once been available for public use (1972-1998) and having formed the sampling frame for the New Immigrant Survey cohorts of 1996 and 2003. And the linchpin procedure is using the unique A-Number identifier to match the basic Immigration File to several other DHS data sets, following the classical matched data sets linking new LPRs to naturalization data, also well-known and once available for public use. The questions on which we focus pertain, separately by cohort, to the nativity of the U.S.

citizen sponsors of spouses, the proportion of conditional LPRs whose conditionality restrictions are removed and the proportion who instead become deportable, the proportion of new LPRs who sponsor the immigration of relatives (spouse; unmarried children) and the number of sponsored relatives, the proportion who acquire citizenship thru naturalization and its timing, the proportion who acquire citizenship by deriving it from a parent and its timing, the proportion who disappear after ten years (viz., because they neither become citizens nor renew their green card), the proportion who sponsor the immigration of relatives after they become citizens and the number of sponsored relatives, and occupational history (five-year employment history at LPR, supplemented at each additional contact with USCIS). The framework envisions a sequence of stages, progressing to enhancements for current information and to additional questions. In this way, a rich new data base can grow.

Discussant: Steve Payson, University of Maryland

Paper #2

Paper Title: Funding Climate Justice: Taxation transfers and green bonds

Presenter: Julia M. Puaschunder, Columbia University

Abstract:
Five indices are empirically created to determine which countries should be using taxation to raise funds for the climate bonds issuance and which countries be granted transfer payments via climate bonds premiums. Index 1 ranks countries based on the country's initial position on a relative climate change economic gains and losses index and the country's CO_2 emissions. Index 2 ranks countries based on the country's initial position on a relative climate change economic gains and losses index and climate flexibility in temperature ranges and its CO_2 emissions. Index 3 ranks countries based on the country's initial position on a relative climate change economic gains and losses index, the country's CO_2 emissions and CO_2 emissions changes. Index 4 ranks countries based on the country's initial position on a relative climate change economic gains and losses index and the country's CO_2 emissions, as well as the estimated lending rate of bonds in that country. Index 5 ranks countries based on the country's initial position on a relative climate change economic gains and losses index and the country's consumption-based trade-adjusted CO_2 emissions. The countries economically gaining from climate change and being climate flexible as well as countries with high CO_2 emissions and not changing CO_2 emissions levels consuming goods and services from other countries as well as having favorable bank lending rates could be taxed for climate bonds issuance to transfer funds via climate bonds to regions of the world that are losing from global warming and are not climate flexible as well as countries with low CO_2 emissions and lowering CO_2 emissions levels that are producing goods and services that are consumed in other parts of the world as well as having unfavorable bank lending rates and hence higher industry financing costs.

Discussant: Marvin Hoekema, USDA/AMS/Dairy Program

Paper #3

Paper Title: Data-driven Investigation on Impact of Air Quality in Different Demographics

Presenter: Areerat Kichkha, AIRLEAP

Additional Authors: Jaelin Lee, Digital Built National Capital Region; Irina Amari, Netsas

Abstract:
This paper investigates air quality to see how levels of air pollution are different in various demographical communities, employing a data science approach, for implications to improve health and well-being equality. Studies have shown that some communities in the United States are disproportionally affected by air pollution levels. This research focuses on the evolution of air quality and main environmental pollutants in different communities in the past few years, including the COVID-19 lockdown period. The lockdown impacted human behaviors, traffic patterns, industrial activities, and commercial and residential building operating schedules. As a result, some communities experienced improvement in air quality, while other communities experienced the opposite in regard to certain pollutants. Our research analyzes air pollutants, weather, and census data provided by the United States Environmental Protection Agency (EPA) to answer key questions such as: 1) Did COVID-19 improve the air quality equally across communities with different demographics? 2) Is economic status a proxy for the likelihood of getting exposed to health-threatening air pollutants? 3) What recommendations can we consider for promoting environmental justice in different communities in the United States?

Discussant: Scott Gilbert, Southern Illinois University Carbondale

---

Paper #4

Paper Title: A Long Short-Term Memory (LSTM) Approach to Forecast the Macroeconomic Impacts of COVID-19 on G-20 Countries

Presenter: Brian W Sloboda (University of Maryland, Global Campus)

Additional Authors: Rolando Santos, Lakeland Community College

Abstract:
The disease COVID-19 initial impacts disrupted the world due to the lockdowns and restrictions on travel. Economic burden rises as the variants continue to persist. This paper explores the use of the deep learning-based algorithms Long Short-Term Memory (LSTM) to assess the economic impacts on the G-20 countries. The LSTM-based models incorporate additional "gates" to memorize longer sequences of input data. Earlier empirical studies conducted in the physical sciences and other disciplines show that the use of the LSTM outperforms traditional forecasting models such as the ARIMA. We used data for the G-20 countries for various macroeconomic variables such as unemployment rate, interest rates,

inflation, retail sales, industrial production, Brent crude oil, and other macroeconomic variables from 1995.1 to 2022.1 to build an LSTM forecasting model. This LSTM model would emphasize the impacts of each macroeconomic variable; e.g., the interest rate policy, and how these values are consistent with the forecasted values. The continued recovery of each G-20 country may be different because of the macroeconomic differences that exist between the G-20 countries.

Discussant: Sandip Sureka, Cornell College

_____