

COEURE's workshop on
“Developments in Data and Methods for Economic Research”

July 2-3, 2015

Université libre de Bruxelles
42, avenue Roosevelt – 1050 Brussels Building SBS-EM
Solvay Brussels School of Economics

Background on COEURE

The **COEURE Coordination Action** (which stands for COoperation for EUropean Research in Economics, www.coeure.eu) brings together the key stakeholders in the European economic research space –policy-makers, researchers, data providers, funders – in a process of stocktaking, consultation and stakeholder collaboration. The aim of this process is to identify the knowledge gaps on key EU economic policy issues and assess the current challenges and opportunities facing European-based researchers, with the ultimate goal to inform the European agenda for research funding. COEURE is funded by the FP7 SSH research programme and it is coordinated by the European Economic Association.

Objectives of the workshop:

The “data and methods” workshop is one of 12 thematic workshops that are being organized to assess the state of the art. Its objective is to explore developments and prospects (1) in the production, funding, access and use of data for economic research and policy, and (2) in the methods of inquiry used in economic research. The workshop will bring together a small group of high level scientists, representatives of statistical offices and other producers of data, research funders as well as policy-makers to discuss these issues, identify good practices and derive recommendations for European research and research infrastructure funding.

Scientific committee

Estelle Cantillon (ULB and COEURE), Christine De Mol (ULB), Bram De Rock (ULB), Domenico Giannone (ULB and Fed Reserve of NY), Georg Kirchsteiger (ULB) and Laszlo Matyas (Central European University and COEURE).

Registration

Registration is free but compulsory. To maximize interactions, attendance is limited to 100 people, on a first-come first-served basis. Contact Nancy De Munck at Nancy.De.Munck@ulb.ac.be to register.

Program

Day 1:

Welcome, description of objectives of COEURE and of the day (9:00-9:15), Estelle Cantillon (ULB) and Marianne Paasi (European Commission, DG RTD)

Session 1: Organizing data access for research (9:15-11:15) (Chair: Estelle Cantillon)

Issues to be addressed: Two key sources of new data on which recent economic research draws are linked individual or firm-level administrative data, and private datasets. Protocols for secure access, de-identifying and confidentiality are essential. What are the keys to the Scandinavian success in making quality administrative data available to researchers? How transposable are these models to other countries, and to other types of confidential data (financial data, private data, ...)? What are the remaining limitations? Is the European legal framework adequate? What are the implications for funding mechanisms?

Speakers: Roberto Barcellan (Eurostat, decentralized remote access to confidential micro-data), Caterina Calsamiglia (CEMFI) and Frederic Udina (Statistical Institute of Catalonia) for the Catalan case / education data, Vigdis Kvalheim (Norway Social Science Data Service), Luke Sibieta (IFS and UK Administrative Data Research Network)

Coffee break

Session 2: Researcher-generated databases (11:40-13:00) (Chair: Bram De Rock)

Issues to be addressed: Partly in response to frustration with cross-country comparison of data or simply lack of centrally collected data, a number of databases have been created and made public by teams of researchers. What are the benefits and costs of such researcher-generated databases? To what extent have they been adopted by other researchers? How are they funded? Should funders impose conditions such as methodologies, quality and open access? How can their survival be organized?

Speakers: Julia Lane (New York University), Arie Kapteyn (University of Southern California), Axel Boersch-Supan (Max Planck Institute, coordinator of SHARE)

Lunch



Session 3: Data generation in (quasi)controlled environments (2:10-3:30) (chair: Georg Kirchsteiger)

Issues to be addressed: These past 20-30 years have seen the development of new datasets generated in controlled or quasi-controlled environments: lab experiments, field experiments, randomized controlled trials, and more recently biomedical (neuro-imaging) data. What have these new types of data allowed us to discover that other more traditional datasets had missed? How robust are the results generated from these data? To what extent can such methods be used for policy design? Can such datasets be useful beyond the specific research project they were designed for? How?

Speakers: Bruno Crépon (CREST and JPAL) and Colin Camerer (Cal Tech)

Coffee break

Session 4: Data standards and cross-country datasets (3:50 – 6:00) (chair: Bram De Rock)

Issues to be addressed : Cross-country variations in data standards and data definition are big obstacles to multi-country research and comparative analyses that are nevertheless crucial for better understanding the scope for replicability of policies across borders (is the experience of country X relevant for country Y?). Are researchers' needs different from other users of such data? What are the main obstacles to greater data comparability? Where does Europe stand relative to the US? Is there need for greater coordination? What are interesting developments in multi-country quality datasets and what are their success drivers? How are they funded? What new insights have such quality multi-country datasets generated (or could generate)?

Speakers: Roberto Barcellan (Eurostat) on the G20 data gap initiative, Peter Bøegh-Nielsen (Statistics Denmark) on two int'l projects for harmonized datasets, Laszlo Halpern (Hungarian Academy of Science and coordinator of MAPCOMPETE), Joseph Tracy (NY Fed) on international banking data, Lisa Wright (Bureau Van Dijk, Zephus) on creating reliable international firm-level merger data.

Day 2

Session 5: The changing face of research-policy / research private sector collaborations (9:00-10:30) (chair: Estelle Cantillon)

Issues to discuss: Researchers have long been involved in policy as ex-post evaluators (policy assessment) or as advisors but new, more collaborative models of interactions are emerging where researchers and policy-makers are partnering, with the benefits of access to data and possibly funding for the former and quality advice for the latter (such partnerships are also developing between researchers and data-intensive firms such as Yahoo, Microsoft, Google, and financial exchanges to name just a few). What are the implications of these developments for the type of research being carried out, or for the organisation of this research? What are the benefits and risks? (e.g. in terms of scientific integrity and independence, data confidentiality



and thus non replicability of the results, ...). What impact is this likely to have for the current “publication as quality certification” model?

Speakers: Liran Einav (Stanford), Asim Khwaja (Harvard) and Markus Moebius (Microsoft)

Coffee break

Session 6: Panel on the implications of the developments in data and methods in economics for research funding (10:45-12:00) (chair: Véronique Halloin, Head of the Belgian French-speaking Research Fund (FNRS))

Participants: Antonio Di Giulio (European Commission, DG RTD), Angelika Kalt (Swiss National Science Foundation) and Paul Sanderson (ESRC, UK)

Lunch

Session 7: Big data: Definition, challenges and opportunities (1- 2 pm) (chair: Christine De Mol)

Issues to be addressed: What is big data, where has it or is it expected to emerge in economics, and how is it likely to change economics? What are examples of big data applications delivering fundamentally new research insights? What are the challenges? How is Europe positioned to take advantage of big data?

Speakers: Sendhil Mullainathan (Harvard) and Lucrezia Reichlin (London Business School)

Coffee break

Session 8: How will big data change econometrics? (2:15-4:15) (chair: Domenico Giannone)

Issues to be addressed: What challenges do high dimensionality data create for econometrics? What developments in methods are needed to address them? What are the implications of big data for model fit and model selection? What progress can be expected in dimensionality reduction? Is there a common ground to be found between the standard econometric approach and the need to let the data speak? What synergies may arise with statistics, computer sciences and other disciplines?

Confirmed speakers: Eric Gautier (Toulouse), Domenico Giannone (NY Fed), Jeff Wooldridge (Michigan State University), Herman van Dijk (Rotterdam)

Farewell

Speakers biographies

Roberto Barcellan is Head of the Unit in charge of Methodology and Corporate Architecture at Eurostat, the statistical office of the European Union. He covered different managerial positions in Eurostat: Head of the Units in charge of Price Statistics, Purchasing Power Parities, Housing Statistics and National Accounts – Production and secretary of the Committee on Monetary, Financial and Balance of Payments Statistics (CMFB). His current portfolio includes methodology, research and innovation, confidentiality and enterprise architecture and the programme for the modernisation of the European Statistical System production processes. He holds a PhD in Statistics from the University of Padua, Italy.

Peter Bøegh-Nielsen, PhD, is head of division for statistics on structural business statistics, including statistics on globalization, at Statistics Denmark. He has been chairing several European development projects on measuring economic globalisation, global value chains and international sourcing. He is currently heading a European project on linking of micro data of business statistics, including the design of a database containing harmonized information at enterprise level from 10 European countries. He is a member of the Bureau of the OECD Working Party on Globalisation of Industry and has been chairman of the bureau of the United Nations city group (Voorburg Group on services statistics) and chairman of OECD's Working Party on Indicators on the Information Society.

Axel Boersch-Supan is director of the Munich Center for the Economics of Aging at the Max Planck Institute for Social Law and Social Policy. His research interests are pensions and retirement behavior, household savings behavior, macroeconomics of population aging and financial markets, and panel data and discrete choice econometrics. He has extensive policy exposure, both nationally and internationally, and was, among others, member of the Council of Advisors to the German Economics Ministry, and co-chaired the German Pension Reform Commission. He coordinates the Survey of Health, Ageing and Retirement in Europe (SHARE).

Caterina Calsamiglia is a research professor of economics at CEMFI and affiliated professor from the Barcelona GSE. Her research focuses on public economics, with an emphasis on school choice, affirmative action and welfare economics. She is the first researcher in Catalunya to get access to linked administrative data on education outcomes and related socio-economic data. She recently received an ERC grant for her work on school choice.

Colin Camerer is the Robert Kirby Professor of Behavioral Finance and Economics at the California Institute of Technology where he teaches cognitive psychology and economics. A pioneer in behavioral economics and neuroeconomics, he is interested in how psychological forces and their deeper neuroscientific foundations influence economic decisions involving individuals and markets. In his research, he uses experiments to better understand how individuals and markets function, neuroscience to gain insight into the neuroscientific drivers for decision making and behavior, and game theory. He is chair of the Russell Sage Foundation Behavioral Economics Roundtable and was named a MacArthur Fellow in 2013.

Bruno Crépon is Professor of economics at ENSAE. He is a board member of the Abdul Latif Jameel Poverty Action Lab (J-PAL), a research network devoted to randomized evaluations of social programs, where he co-chairs the Employment Program. Crépon is also an IZA and CEPR Research Associate. Crépon completed his undergraduate studies at Ecole Polytechnique in Paris in 1986 and ENSAE in 1988. He completed his PhD in Economics at Université de Paris I in 1994. His research focuses on program evaluation, especially on employment and youth employment programs in both developed and developing countries. He has conducted many randomized evaluations in France, Morocco, Egypt, Cote d'Ivoire and South Africa.

Antonio Di Giulio is head of unit "Research Infrastructures" at DG Research of the European Commission. He holds a Masters degree in agricultural sciences and a specialised Masters degree in rural development, programmes and projects. He has a Doctoral degree in food economics, and is an elected member of the Accademia dei Georgofili, Florence (Italy).

Liran Einav is a Professor of Economics at Stanford University and a Research Associate in the National Bureau of Economic Research. Einav's areas of specialization are industrial organization and applied microeconomics. An important strand of his



work is focused on health and other insurance markets, including the development of empirical models of insurance demand and pricing, and empirical analyses of the implications of adverse selection and moral hazard. Einav has also studied credit markets and more recently online markets, often making use of large proprietary datasets. He is co-editor of *Econometrica*.

Eric Gautier is professor of econometrics at the University of Toulouse I. His research interests include statistics and econometrics with high dimensional parameters.

Domenico Giannone is Senior Economist at the Federal Reserve Bank of New York and Research Fellow of the Centre for European Policy Research (CEPR). His general fields of research are forecasting, monetary policy, business cycles and growth. He has designed econometric models that are routinely used to inform policy decisions in many institutions, was co-founder and director of Now-Casting.com (a web-based forecasting company), and was a member of the CEPR Business Cycle Dating Committee. He is associate editor for the *Journal of Applied Econometrics*, the *International Journal of Forecasting* and *Empirical Economics*.

Véronique Halloin is Secretary General of the Belgian French-speaking Research Fund (FNRS) whose goal is to promote fundamental research in the French-speaking Community of Belgium. She holds a Ph.D. in Civil Engineering and was a full professor of Chemical Engineering at the Université Libre de Bruxelles and Vice-Rector for Research (2006-08) prior to her current appointment. As Secretary General of the FNRS, she also represents Belgium in various international research and research-related bodies such as the OECD's Global Science Forum, CERN, and the European Foundation for Science.

László Halpern is director at the Centre for Economic and Regional Studies of the Hungarian Academy of Sciences in Budapest and a research fellow in the International Macroeconomics and International Trade research programmes of the Centre for Economic Policy Research (London). He has written widely on exchange rate and exchange rate policy in Central and Eastern Europe; on enterprise behaviour, microeconomic environment and economic policy; on multinational enterprises, foreign direct investment and economic development. He coordinates the FP7 project MAPCOMPETE whose objectives are to analyze data opportunities and requirements to analyze and compare competitiveness in EU countries.

Angelika Kalt is Deputy Director of the Swiss National Science Foundation, where she is responsible for quality and development issues concerning research funding and assessment. Dr. Kalt is also the head of the SNSF division that promotes interdisciplinary and cooperative research. She was previously professor in Earth Sciences (petrology and geodynamics) at the University of Neuchâtel.

Arie Kapteyn is professor of economics and the founding Executive Director of the Dornsife Center for Economic and Social Research at the University of Southern California. Much of Professor Kapteyn's recent research is in the field of aging and economic decision making, with papers on topics related to retirement, consumption and savings, pensions and Social Security, disability, economic well-being of the elderly, and portfolio choice. He is the founder of several internet panels including the CentERpanel at Tilburg University, the American Life Panel, a nationally representative sample of 6,000 households maintained at RAND, and the Understanding America Study (2,000 households) at USC.

Asim Khwaja is the Sumitomo-Foundation for Advanced Studies on International Development Professor of International Finance and Development at the Harvard Kennedy School, and Co-Director of Evidence for Policy Design (EPoD). His areas of interest include economic development, finance, education, political economy, institutions, and contract theory/mechanism design. His research combines extensive fieldwork, rigorous empirical analysis, and microeconomic theory to answer questions that are motivated by and engage with policy.

Vigdis Kvalheim is Deputy Director of the Norwegian Social Science Data Services (NSD) and manages, among others, the Data Protection Services and the Individual Level Data Unit at NSD. She is or has been involved in several Norwegian, Nordic and international data management activities, EU projects, working groups and advisory boards. She is a member of the Board of Directors of the International Federation of Data Organizations for the Social Sciences (IFDO), the NordForsk's high-level group on research infrastructures and chairs the Project Board for the Norwegian Remote Access Infrastructure for Register Data (RAIRD). She was, until recently acting director of CESSDA, the umbrella organization for the European national data archives.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under the grant agreement nr 320300



Julia Lane is Professor of Public Policy at the Wagner School and Professor of Practice at the Center for Urban Science and Progress at New York University. She has carried out a number of data-driven projects over her career. She is, among others, the founder of the Longitudinal Employer-Household Dynamics Program at the US Census Bureau (which evolved into a permanent Census Bureau Program); she has created and developed the NORC/University of Chicago data enclave to provide remote and protected access to sensitive microdata and has overseen a number of social surveys in the US.

Markus Mobius is principal researcher at Microsoft Research New England. His primary research interests include the economics of social networks where he seeks to develop models of learning, coordination and cooperation in social networks and to test those using field and lab experiments. His recent work explores how browsing data can be leveraged to analyze news consumption.

Sendhil Mullainathan is Professor of Economics at Harvard University. His research interests span a broad spectrum including behavioral economics, labor economics, the market for media and corporate finance. His latest research focuses on using machine learning and data mining techniques to better understand human behavior. He helped co-found a non-profit to apply behavioral science (ideas42), co-founded a center to promote the use of randomized control trials in development (the Abdul Latif Jameel Poverty Action Lab), and has worked in government in various roles, including most recently as Assistant Director of Research at the Consumer Financial Protection Bureau.

Lucrezia Reichlin is Professor of Economics at the London Business School, co-founder and director of Now-Casting Economics Ltd, non-executive director of UniCredit Banking Group and AGEAS Insurance Group. She is Chair of the Scientific Council at the Brussels based think-tank Bruegel as well as a member of the Commission Economique de la Nation (advisory board to the French finance and economics ministers). Between March 2005 and September 2008 she served as Director General of Research at the European Central Bank. She is an expert on forecasting, business cycle analysis and monetary policy. The econometric methods she has developed for short term forecasting (now-casting) are widely used in central banks around the world.

Paul Sanderson leads on economics and finance at the UK Economic and Social Research Council (ESRC), where he is responsible for the development of the ESRC's research portfolio in these areas and for developing academic engagement and collaboration with the financial services sector. Paul has been with the ESRC since 2010. He began his career as an academic economist but has subsequently held positions in management and research at the Bank of England, HM Treasury and with a number of firms in the UK private sector.

Luke Sibieta is programme director of the education and skills sector at the Institute for Fiscal Studies. He has wide experience in accessing and conducting research using administrative data. He is also a non-executive board member of the UK's new Administrative Data Research Network (<http://www.adrn.ac.uk/>), an initiative that is seeking to make more administrative and linked data available to researchers (funded by the Economic and Social Research Council).

Joseph Tracy is Executive Vice President and Senior Advisor to the President at the Federal Reserve Bank of New York. He is, among others, a member of the Global Legal Identifier Project to create a global system of unique identifiers for legal entities and chair of Data Workstream for the Data Gaps Implementation Project sponsored by the Financial Stability Board. Tracy holds a Ph.D. in Economics from the University of Chicago. His research interests include unions and collective bargaining as well as housing and urban economics.

Frederic Udina is the general manager at the Statistics Bureau of Catalonia (Idescat) since 2011 and president of the Catalan Institute of Public Policy Evaluation (Ivàlua) since 2013. Frederic is also professor in math, probability and statistics at Universitat Pompeu Fabra and at the Barcelona Graduate School of Economics. He holds a PhD in Mathematics from the Universitat Politècnica de Catalunya and a degree in Mathematics from the Universitat Autònoma de Barcelona. He has published numerous articles, has participated in multiple international research groups and performed consulting in statistics for Morgan Stanley, among others.

Herman van Dijk is professor of econometrics at VU University Amsterdam and professor emeritus at Erasmus University Rotterdam where he was director of the Tinbergen Institute and director of the Econometric Institute. He is the co-founder of

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under the grant agreement nr 320300



the EC2 meetings of European econometricians; cofounder of the European Seminar on Bayesian Econometrics (ESOB) and cofounder of the Econometric and Tinbergen Institute lectures that are published by Princeton University Press. His research interests cover a range of topics in econometrics with, as common themes: Simulation-based Bayesian Econometric Techniques for Inference, Forecasting and Decision analysis. His recent research involves econometric forecasting with large data sets and advanced algorithms using parallel computing.

Jeffrey Wooldridge is University Distinguished Professor of Economics at Michigan State University where he has taught since 1991. His research interests include econometrics of cross section and panel datasets. He is a Fellow of the Econometrics Society, of the Journal of Econometrics and a recipient of the Alfred P. Sloan Fellowship.

Lisa Wright is managing director at Bureau Van Dijk, a publisher of global business information (ORBIS, Bankscope, Zephyr). She is responsible for BvD's M&A products globally and manages Zephyr Ltd, the BvD subsidiary that specializes in researching and creating their Global M&A, PE, IPO & Venture Capital deals database, Zephyr.