

Christine Choirat

Curriculum Vitae

August 2019

Address: Swiss Data Science Center, INN Building, Station 14, 1015 Lausanne, Switzerland

Phone: +41 79 291 76 42

Email: cchoirat@gmail.com, cchoirat@datascience.ch

WWW: <http://scholar.harvard.edu/cchoirat/>

Skype: christine_choirat

GitHub: <https://github.com/cchoirat/>

Research expertise

Health data science and analytics
Data science platforms and workflows
Reproducible research
High-performance computing
Statistical software development
Environmental policy, health policy
Causal inference methods

Employment

Chief Innovation Officer

2019-present

Swiss Data Science Center (SDSC), an ETH Domain initiative for Data Science in Switzerland

- Work with the Executive and Academic Directors on long-range strategic planning to ensure efforts and research directions are consistent with the goals of the SDSC
- Create, plan, develop, and facilitate relationships between the SDSC and academic, governmental and intergovernmental organizations, nationally and internationally
- Promote the data science platforms, tools and services offered at the SDSC
- Foster interdisciplinary collaboration and research in Data Science, Health IT, and Open Science
- Supervise and mentor junior and senior data scientists, as well as Master and PhD students, and postdoctoral fellows

2019 (Jul-present): Chief Innovation Officer

2019 (Jan-Jun): Chief Health Data Scientist

Adjunct Lecturer

2017-present

Swiss Federal Institute of Technology Lausanne (EPFL), 2019-present

- Instructor for the course "EE-490(h): Lab in Data Science", a course in the Masters of Science in Data Science

Department of Biostatistics, Harvard T.H. Chan School of Public Health, 2017-2018

- Instructor for the course "BST262: Computing for Big Data", a course in the Masters of Science in Health Data Science
- HarvardX module on computational tools for reproducible research (<https://www.edx.org/course/principles-statistical-computational-harvardx-ph527x>)

Harvard Extension School, 2019-present

- Instructor for the course "CSCI E-88B: Computing for Big Data", a course in the Master of Liberal Arts, Data Science

2019: Adjunct Lecturer

2017-2019 : Instructor

Senior Research Scientist

2015-2018

Department of Biostatistics, Harvard T.H. Chan School of Public Health

- Lead data scientist (National Studies on Air Pollution and Health, PI's: F. Dominici, A. Zanobetti, C. Zigler)
- Lead the effort of implementing software solutions to assess effectiveness of regulatory measures via advanced statistical methods
- Advise PIs on best practice in data science, software development and high-performance computing
- Supervise or co-supervise over 15 junior data scientists, PhD students and postdocs
- Contribute to scientific articles and grant proposals

- Teach masters, doctoral and postdoctoral courses on big data analytics
- Serve on search committees for data scientists, developers and postdoctoral students
 - 2018: Senior Research Scientist
 - 2015: Research Scientist

Research Associate

2014-2015

Institute for Quantitative Social Science, Harvard University

- Core Team Member of the suite of R packages Zelig (<http://zeligproject.org/>): architecture redesign, community building, real-time documentation
- Core developer of genTB (<https://gentb.hms.harvard.edu/>), a web platform for translational genomics of tuberculosis: genotype-based predictions of drug resistance, mapping of genetic and phenotypic data, data sharing and exploration

Invited Visiting Scholar

April-August 2013

Institute for Quantitative Social Science, Harvard University

Invited Visiting Professor

June 2012

Department of Economics and Finance, LUISS University, Rome, Italy

Director of Research

2011-2013

School of Economics and Business Administration, University of Navarra, Pamplona, Spain

- Defined School research priorities, coordinated research groups, organized research seminars and promoted the research production of over 40 faculty members
- Strengthened industry partnerships
- Chaired workshops and conferences
- Served on and led search committees for faculty

Professor of Quantitative Methods

2006-2013

School of Economics and Business Administration, University of Navarra, Pamplona, Spain

- Designed and taught a broad portfolio of quantitative courses at the undergraduate and graduate levels
- Aided student recruitment
- Coordinated student international exchange programs.
 - 2013: Professor (tenured)
 - 2009-2013: Associate professor (tenured)
 - 2006-2009: Assistant professor (tenure-track)

Education and awards

2009–2012: Best paper awards (2009, 2010, 2011, 2012), School of Economics and Business Administration, University of Navarra, Pamplona, Spain.

2008: Best young researcher award, School of Economics and Business Administration, University of Navarra, Pamplona, Spain.

2004-2006: Postdoctoral fellow, Department of Economics, University of Insubria, Varese, Italy.

1999–2003: PhD in Applied Mathematics (*Summa Cum Laude*, major in Statistics), Paris Dauphine University, Paris, France. Funded by the French Ministry of Higher Education.

1998–1999: Master's Degree in Mathematical Finance, Paris Dauphine University & ENSAE, Paris, France. Funded by the French Ministry of Higher Education.

Software

Zelig (<http://zeligproject.org/>)

Unified interface to wide ranging statistical models, that automates graphics, and translates statistical estimates into interpretable quantities of interest

genTB (<https://gentb.hms.harvard.edu/>)

Genomics analysis collaborative web application for tuberculosis data exploration

Grants

- R01ES028033 (Laden) - NIH** **12/15/2017 - 11/30/2022**
Relationship Between Multiple Environmental Exposures and CVD Incidence and Survival: Vulnerability and Susceptibility
 Role: Co-Investigator
- 4953-RFA14-3/16-4 (Dominici/Zanobetti) - HEI** **02/01/2016 - 01/31/2020**
Assessing adverse health effects of long-term exposure to low levels of ambient air pollution
 Role: Co-Investigator
- R01ES026217 (Zigler) - NIH** **02/01/2016 - 01/31/2021**
Causal Inference with Interference for Evaluating Air Quality Policies
 Role: Co-Investigator
- P01 CA134294 (Lin) - NIH** **09/05/2013 - 06/30/18**
Statistical Informatics for Cancer Research
 Role: Research Associate
- R01 ES024332 (Zanobetti) - NIH** **05/01/2015 - 02/28/2019**
Cardiovascular Health and Air Pollution: A National Study
 Role: Research Associate
- 83587201-0 (Koutrakis) - EPA grant (ACE Center)** **12/01/2015 - 11/30/2020**
Regional Air Pollution Mixtures: The Past and Future Impacts of Emission Controls and Climate Change on Air Quality and Health
 Role: Co-Investigator

Academic service

Department

Computing Committee, Department of Biostatistics, Harvard T.H. Chan School of Public Health (2015-2018)

Executive Committee, Master of Science in Health Data Science, Department of Biostatistics, Harvard T.H. Chan School of Public Health (2016-2018)

Guest editor

Special issue of *Business & Society* (2019)

Special issue of *The Spanish Journal of Psychology* (2013)

Conferences

Member of the scientific committee of useR! 2019

Member of the scientific and organizing committee of Harvard DataFest 2017 and 2018 (a Data Science Bootcamp for better research)

Member of the program committee of SciPy 2014, SciPy 2015, SciPy2017, SciPy 2018 and SciPy 2019 (Scientific Computing with Python)

Member of the scientific and organizing committees of EBEN RC 2013 (European Business Ethics Network Research Conference)

Chair of EMPG 2012 (41st European Mathematical Psychology Group Meeting)

Referee

Reviewer for the National Science Foundation, U.S.A.

Expert for the "Agencia Nacional de Evaluación y Prospectiva" (ANEP), Spain

Ad-hoc reviewer (*Omega*, *Econometric Theory*, *Scientific Reports*, *Scandinavian Journal of Psychology*, *Scientific Data*)

Teaching experience

PhD Statistical computing with R, Data science with R

Masters and MBA Computing for big data, Microeconometrics, Financial econometrics, Mathematical methods for management, Computational statistics

Undergraduate Statistics, Optimization and linear programming, Calculus, Microeconometrics

Languages

Native: French

Full Professional Proficiency: English, Spanish, Italian

Basic German

Recent invited talks and conferences

Keynote session chair, useR! 2019 (annual R Foundation conference), Toulouse, France, July 2019

e-IRG Workshop (e-Infrastructure Reflection Group, e-IRG is a strategic body to facilitate integration in the area of European e-Infrastructures and connected services, within and between member states, at the European level and globally), CERN, Switzerland, May 2019

Environmental Statistics Seminar, Harvard T.H. Chan School of Public Health, December 2018

Department of Mathematical Sciences, United States Military Academy, West Point, October 2018

Publications

Refereed journal articles

- [1] C. Choirat, C. Hess, and R. Seri. A functional version of the Birkhoff ergodic theorem for a normal integrand: A variational approach. *The Annals of Probability*, 31(1):63 – 92, 2003.
- [2] M. Bernasconi, C. Choirat, and R. Seri. Measurement by subjective estimation: Testing for separable representations. *Journal of Mathematical Psychology*, 52(3):184 – 201, 2008.
- [3] C. Choirat and R. Seri. Statistical properties of generalized discrepancies. *Mathematics of Computation*, 77(261):421 – 446, 2008.
- [4] C. Choirat and R. Seri. Econometrics with Python. *Journal of Applied Econometrics*, 24(4):698–704, 2009.
- [5] M. Bernasconi, C. Choirat, and R. Seri. The analytic hierarchy process and the theory of measurement. *Management Science*, 56(4):699 – 711, 2010.
- [6] C. Hess, C. Choirat, and R. Seri. Ergodic theorems for extended real-valued random variables. *Stochastic Processes and their Applications*, 120(10):1908 – 1919, 2010.
- [7] M. Bernasconi, C. Choirat, and R. Seri. A re-examination of the algebraic properties of the AHP as a ratio-scaling technique. *Journal of Mathematical Psychology*, 55(2):152 — 165, 2011.
- [8] C. Choirat and R. Seri. Estimation in discrete parameter models. *Statistical Science*, 27(2):278 – 293, 2012.
- [9] C. Choirat and R. Seri. Computational aspects of Cui-Freeden statistics for equidistribution on the sphere. *Mathematics of Computation*, 82(284):2137 – 2156, 2013.
- [10] C. Choirat and R. Seri. Numerical properties of generalized discrepancies on spheres of arbitrary dimension. *Journal of Complexity*, 29(2):216 – 235, 2013.
- [11] R. Seri and C. Choirat. Scenario approximation of robust and chance-constrained programs. *Journal of Optimization Theory and Applications*, 158(2):590 – 614, 2013.
- [12] M. Bernasconi, C. Choirat, and R. Seri. Empirical properties of group preference aggregation methods employed in AHP: Theory and evidence. *European Journal of Operational Research*, 232(3):584–592, 2014.

- [13] L. Brandolini, C. Choirat, L. Colzani, G. Gigante, R. Seri, and G. Travaglini. Quadrature rules and distribution of points on manifolds. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, 5:889 – 923, 2014.
- [14] C. Choirat and R. Seri. Bootstrap confidence sets for the Aumann mean of a random closed set. *Computational Statistics & Data Analysis*, 71(0):803 – 817, 2014.
- [15] C. Hess, R. Seri, and C. Choirat. Essential intersection and approximation results for robust optimization. *Journal of Nonlinear and Convex Analysis*, 15(5):979 – 1002, 2014.
- [16] R. Seri and C. Choirat. Comparison of approximations for compound Poisson processes. *ASTIN Bulletin*, 45(3):601 – 637, 2015.
- [17] Q. Di, Y Wang, A. Zanobetti, Y Wang, C. Choirat, J.D. Schwartz, and F. Dominici. Association of short-term exposure to air pollution with mortality in older adults. *Journal of the American Medical Association*, 318(24):2446–2456, 2017.
- [18] Q. Di, Y. Wang, A. Zanobetti, Y. Wang, P. Koutrakis, C. Choirat, F. Dominici, and J.D. Schwartz. Air Pollution and Mortality in the Medicare Population. *The New England Journal of Medicine*, 376(26):2513–2522, 2017.
- [19] J.H. Wasfy, C.M. Zigler, C. Choirat, Y. Wang, F. Dominici, and R.W. Yeh. Readmission rates after passage of the Hospital Readmissions Reduction Program: A pre–post analysis. *Annals of Internal Medicine*, 166(5):324–331, 2017.
- [20] C.M. Zigler, C. Choirat, and F. Dominici. Impact of National Ambient Air Quality Standards Nonattainment Designations on Particulate Pollution and Health. *Epidemiology*, 29(2):165–174, 2018.
- [21] C. Choirat, D. Braun, and M.-A. Kioumourtzoglou. Data Science in Environmental Health Research. *Current Epidemiology Reports*, pages 1–9, 2019.
- [22] R. Chun, A. Argandoña, C. Choirat, and D.S. Siegel. Corporate reputation: Being good and looking good. *Business & Society*, 58(6):1132–1142, 2019.
- [23] Q. Di, H. Amini, L. Shi, I. Kloog, R. Silvern, J. Kelly, M.B. Sabath, C. Choirat, P. Koutrakis, A. Lya-pustin, Y. Wang, L.J. Mickley, and J. Schwartz. An ensemble-based model of PM_{2.5} concentration across the contiguous United States with high spatiotemporal resolution. *Environment International*, 130(June):104909, 2019.
- [24] L.R. Henneman, C. Choirat, C Ivey, K. Cummiskey, and C.M. Zigler. Characterizing population exposure to coal emissions sources in the united states using the HyADS model. *Atmospheric Environment*, pages 271–280, 2019.
- [25] L.R.F. Henneman, C. Choirat, C. Ivey, K. Cummiskey, and C.M. Zigler. Characterizing population exposure to coal emissions sources in the united states using the hyads model. *Atmospheric environment*, 203:271–280, 2019.
- [26] L.R.F. Henneman, C. Choirat, and C.M. Zigler. Accountability assessment of health improvements in the united states associated with reduced coal emissions between 2005 and 2012. *Epidemiology*, 30(4):477–485, 2019.
- [27] G. Papadogeorgou, C. Choirat, and C.M. Zigler. Adjusting for unmeasured spatial confounding with distance adjusted propensity score matching. *Biostatistics*, 20(2):256–272, 2019.
- [28] J. H Wasfy, V. Bhambhani, E. Healy, C. Choirat, F. Dominici, C. Shen, Y. Wang, and R. Yeh. Relative effects of the hospital readmissions reduction program for congestive heart failure on hospitals that serve poorer patients. *Journal of the American College of Cardiology*, 73(9 Supplement 1):729, 2019.

- [29] X. Wu, D. Braun, M.-A. Kioumourtzoglou, C. Choirat, Q. Di, and F. Dominici. Causal inference in the context of an error prone exposure: air pollution and mortality. *The Annals of Applied Statistics*, 13(1):520–547, 2019.
- [30] C. Kim, M.J. Daniels, J.W. Hogan, C. Choirat, and C.M. Zigler. Bayesian methods for multiple mediators: Principal stratification and causal mediation analysis of power plant emission controls. *Annals of Applied Statistics*, Forthcoming.

Articles in progress

- [1] C. Choirat, D Di Cagno, and A. Galliera. Gamblers or investors? An experiment on the almost winning outcomes. In preparation.
- [2] C. Choirat, J. Honaker, M Idris, K. Imai, G. King, O. Lau, and J. McGrath. Interpreting Zelig: Everyone’s statistical software. In preparation.
- [3] K. Cummiskey, C. Kim, C. Choirat, L. Henneman, J.D. Schwartz, and C.M. Zigler. A source-oriented approach to coal power plant health effects. In preparation.
- [4] C. Kim, C. Choirat, and C.M. Zigler. Health effects of power plant emissions mediated through ambient PM. In preparation.
- [5] Y. Wei, Y. Wang, Q. Di, C. Choirat, Y. Wang, P. Koutrakis, A. Zanobetti, F. Dominici, and J.D. Schwartz. Air pollution and cause-specific risks and costs of hospital admissions. In revision.
- [6] G. Papadogeorgou, C.M. Zigler, F. Mealli, F. Dominici, J.H. Wasfy, and C. Choirat. Causal impact of the Hospital Readmissions Reduction Program on hospital readmissions. Submitted.

Refereed book chapters

- [1] C. Choirat and R. Seri. The asymptotic distribution of quadratic discrepancies. In H. Niederreiter and D. Talay, editors, *Monte Carlo and Quasi-Monte Carlo 2004*, pages 61 – 76. Springer Verlag, 2006.
- [2] C Hess, C. Choirat, and R. Seri. Approximation of stochastic programming problems. In H. Niederreiter and D. Talay, editors, *Monte Carlo and Quasi-Monte Carlo 2004*, pages 45 – 60. Springer Verlag, 2006.
- [3] C.M. Zigler, C. Kim, C. Choirat, J.B. Hansen, Y. Wang, L. Hund, J. Samet, G. King, and F. Dominici. Causal inference methods for estimating long-term health effects of air quality regulations. *Health Effects Institute, Research Report 187*, 2016.

Refereed conference proceedings

- [1] C. Choirat and R. Seri. Confidence sets for the aumann mean of a random closed set. In *Proceedings 8th International Conference IPMU*, volume 1, pages 509–514. CSIC (Consejo Superior de Investigaciones Científicas), Madrid, 2000.
- [2] R. Seri and C. Choirat. Stochastic boundedness in biological models. In V. Capasso, editor, *Mathematical Modeling & Computing in Biology and Medicine, 5th ESMTB Conference*, pages 310 – 316. Società Editrice Esculapio, Bologna, 2003.
- [3] C. Choirat and R. Seri. Analytic hierarchy process, a psychometric approach. In J. Andrysek, M. Karny, and J. Kracik, editors, *Multiple Participant Decision Making*, volume 9 of *International Series on Advanced Intelligence*, pages 173 – 179. Advanced Knowledge International Pty, Ltd., 2004.
- [4] C. Choirat and R. Seri. Confidence sets for the Aumann mean of a random closed set. In Antonio Laganà, Marina L. Gavrilova, Vipin Kumar, Youngsong Mun, Chih Jeng Kenneth Tan, and Osvaldo Gervasi, editors, *ICCSA (3)*, volume 3045 of *Lecture Notes in Computer Science*, pages 298 – 307. Springer, 2004.
- [5] C. Choirat and R. Seri. Statistical properties of generalized discrepancies and related quantities. In *Proceedings 10th International Conference IPMU*, volume 1, pages 381–384. Casa Editrice Università La Sapienza, Roma, 2004.

- [6] M. Bernasconi, C. Choirat, and R. Seri. Estimation of separable representations in psychophysical experiments. In L. Barsalou B.G. Bara and M. Bucciarelli, editors, *Proceedings of CogSci 2005, XXVII Annual Meeting of the Cognitive Science Society, Stresa (I)*, pages 244–249. Lawrence Erlbaum Associates, 2005.
- [7] C. Choirat and R. Seri. Approximation of the asymptotic distribution of quadratic discrepancies. In *Atti del Convegno SER2006, Convegno Nazionale delle Ricerche sulle Serie Temporal, Villa Mondragone, Monte Porzio Catone (I)*, 2006.
- [8] C. Choirat and R. Seri. Consistency in conditional volatility models. In *Atti del Convegno SER2006, Convegno Nazionale delle Ricerche sulle Serie Temporal, Villa Mondragone, Monte Porzio Catone (I)*, 2006.
- [9] R. Seri and C. Choirat. Generalized discrepancies on the sphere. In *PAMM-Proceedings in Applied Mathematics and Mechanics*, volume 7, pages 2020039–2020040, 2007.
- [10] C. Choirat and R. Seri. On embedding gretl in a python module. In M.V. Esteban I. Díaz-Emparanza, P. Mariel, editor, *Econometrics with gretl. Proceedings of First Gretl Conference*, pages 219–227. Euskal Herriko Unibertsitatea, 2009.
- [11] M. Bernasconi, C. Choirat, and R. Seri. Differentials of eigenvalues and eigenvectors in undamped discrete systems under alternative normalizations. In *Proceedings of WCE2011, World Congress of Engineering (track 2011 International Conference of Applied and Engineering Mathematics), London (UK)*, 2011.
- [12] C. Choirat and R. Seri. A comparison of approximations for compound Poisson processes. In *Proceedings of the 58th ISI World Statistics Congress, Dublin (IRL)*, 2011.
- [13] C. Choirat and R. Seri. Computational aspects of discrepancies for equidistribution on the hypercube. In *Proceedings of COMPSTAT 2012, Limassol (CY)*, 2012.
- [14] R. Seri and C. Choirat. Comparison of quadrature rules for the Wielandt-Nyström method with statistical applications. In *AIP Conf. Proc. 1479, Numerical Analysis and Applied Mathematics ICNAAM 2012: International Conference of Numerical Analysis and Applied Mathematics*. American Institute of Physics, 2012.
- [15] M.B. Sabath, Q. Di, D. Braun, F. Dominici, and C. Choirat. airpred: A flexible R package implementing methods for predicting air pollution. In *IEEE International Conference on Data Science and Advanced Analytics (DSAA) online*, 2018.