BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Stern, Hal Steven

eRA COMMONS USER NAME (credential, e.g., agency login): halstern

POSITION TITLE: Chancellor's Professor, Department of Statistics

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
Massachusetts Institute of Technology	B.S.	06/1981	Mathematics
Stanford University	M.S.	04/1985	Statistics
Stanford University	Ph.D.	08/1987	Statistics

A. Personal Statement

I am Chancellor's Professor in the Department of Statistics at the University of California, Irvine, and until recently, Dean, Donald Bren School of Information and Computer Sciences. I have a long-standing research program in developing and studying Bayesian statistical methods in the context of applications in the biological/health, physical and social sciences. Examples include spatial models for disease incidence and genomic data, analysis of reliability and reproducibility of imaging studies, and statistical methods for missing data in longitudinal data analyses. I also have extensive experience in a wide range of collaborative biostatistics research projects. I have previously served as director of the Biostatistics, Epidemiology and Research Design (BERD) unit of the UCI Institute for Clinical and Translational Science and as co-chair of the Statistics Working Group of the NIH-funded multi-campus brain imaging FBIRN grant. I currently collaborate with neurologists and psychologists as PI of the Biostatistics, Computation and Data Management Core of the Conte Center at UC Irvine funded by NIMH to measure fragmentation and unpredictability of maternal and environmental signals on the developing brain and assess the impact of these patterns on multiple mental-health-relate outcomes. The proposed Conte Center renewal builds on this effort to test the hypothesis that fragmentation and unpredictability of early life maternal and environmental signals (FRAG) contribute to adolescent vulnerabilities and adult mental illnesses via mechanisms involving aberrant development and maturation of emotional brain circuits.

B. Positions and Honors

Positions and Employment

1987-1991 Assistant Professor, Department of Statistics, Harvard University, Cambridge, MA
1991-1994 Associate Professor, Department of Statistics, Harvard University, Cambridge, MA
1994-1997 Associate Professor, Department of Statistics, Iowa State University, Ames, IA
1997-2002 Professor, Department of Statistics, Iowa State University, Ames, IA
2000-2002 Interim Director, L.H. Baker Center for Bioinform. & Biol. Stat., Iowa State University., Ames, IA
2002-2010 Professor and Founding Chair, Department of Statistics, University of California, Irvine, CA
2010-2016 Ted and Janice Smith Family Foundation Dean and Professor of Statistics, Donald Bren School of Information and Computer Sciences, University of California, Irvine, CA
2017- Professor, Department of Statistics, University of California, Irvine, CA
2018- Chancellor's Professor, Department of Statistics, University of California, Irvine, CA

Other Professional Activities

1999-2001	Editor, *Chance* (publication of the American Statistical Association)
2004	Chair, Section on Bayesian Statistical Science, American Statistical Association
2007-2008	Chair, National Academy of Sciences Panel on American Community Survey Use for the NSF Survey of College Graduates
2008-2014	Member, Committee on National Statistics (CNSTAT) of The National Academies
2009-2010	Member, National Academy of Sciences Panel on Missing Data in Clinical Trials
2010	Chair, National Academy of Sciences Steering Committee for a Workshop on the Future of Federal Household Surveys
2010-2012	Editor, Applications & Case Studies (and Coordinating Editor)*, Journal of the American Statistical Association*
2013-2016	Chair, National Academy of Sciences Panel on Research Methodologies for Understanding Commercial Vehicle Driver Fatigue
2014-	Member, Scientific Area Committee for Physics/Pattern Forensic Evidence, Organization of Scientific Area Committees, National Institute of Standards and Technology (NIST)
2015-2017	Chair, American Statistical Association Committee on Publications
2015-	Member, Board of Trustees (Executive Committee), National Institute of Statistical Sciences
2017-	Chair, Section U (Statistics) of the American Association of the Advancement of Science (chair-elect 2017; chair 2018; retiring chair 2019)

Honors and Awards

1983	Alan Abrams Scholarship, Stanford University
1996	Letter of Instructional Commendation, Technical Education Program, GM
1998	Fellow, American Statistical Association
2001-2002	Laurence H. Baker Chair in Biological Statistics, Iowa State University
2001	Buckingham Scholar-In-Residence, Miami University, Oxford, OH
2007	Teaching Excellence Award (Bren School of ICS), University of California, Irvine
2011	Fellow, Institute of Mathematical Statistics
2016	Fellow, American Association for the Advancement of Science
2018	Chancellor's Professor, University of California, Irvine

C. Contribution to Science

1. **Bayesian statistical methods** – My primary research contributions within the field of statistics concerns the development of Bayesian statistical methods and their application in a range of scientific domains. The Bayesian approach is characterized by the use of probability to describe the distribution of unknown quantities and then updating these distributions based on observed information. I am co-author on a leading Bayesian text (the first reference below). Applications in health and technology have included applications to addressing missing data, mapping of disease incidence rates, analysis of genome-wide association studies, and analysis of behavioral data to characterize fragmentation. These applications were carried out with my statistics graduate students and with scientific collaborators.
    a. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. and Rubin, D. B., (2013), *Bayesian Data Analysis*, 3rd edition, Chapman and Hall/CRC: Boca Raton.
    b. Reber, D. L., Stern, H. S., and Berger, P. J. (2000). Comparing traditional and Bayesian analyses of selection experiments in animal breeding. Journal of Agricultural, Biological, and Environmental Statistics, 5, 240-256.
    c. Wright, D., Stern, H. S., and Cressie, N. (2003). Loss functions for estimation of extrema with an application to disease mapping. Canadian Journal of Statistics, 31, 251-266.
    d. Heins, KA and Stern, HS. A statistical model for event sequence data. Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics (AISTATS 2014), 2014; 33:338-346. PMC5397901.

2. **Statistical model checking and model comparison** – It is a critical aspect of data analysis to assess whether an assumed statistical model is appropriate for a given set of data. It is unfortunately quite common for individuals to fit a model and report results without performing such a model assessment. With co-authors and graduate students I have played a key role in developing an approach that can be easily applied in Bayesian data analyses known as posterior predictive model checks. These checks compare the observed data to data sets that might have been generated under the assumed statistical model. Differences between the observed and posterior predictive data suggest a problem with the statistical model.
    a. Gelman, A., Meng, X.-L., and Stern, H. S. (1996). Posterior predictive assessment of model fitness via realized discrepancies (with discussion). *Statistica Sinica*, 6, 733-807.
    b. Stern, H. S. and Cressie, N. (2000). Posterior predictive model checks for disease mapping models. *Statistics in Medicine*, 19, 2377-2397.
    c. Sinharay, S., and Stern H. S. (2003). Posterior predictive model checking in hierarchical models. *Journal of Statistical Planning and Inference*, 111, 209-221.
    d. Sinharay, S. and Stern, H. S. (2005). An empirical comparison of methods for computing Bayes factors in generalized linear mixed models. *Journal of Computational and Graphical Statistics*, 14(2), 415-435.

3. **Statistical analyses in brain imaging studies** – From 2004 through 2011 I was the co-leader of the Statistics Working Group for the Functional Imaging Biomedical Informatics Research Network (FBIRN), a NIH-funded project that aimed to develop methods for carrying out multi-center fMRI studies in the context of schizophrenia. Contributions include the development of novel modeling approaches (e.g., publications b. and d. below) as well as critical assessments of the reliability of fMRI activation in multi-center studies. The FBIRN project has produced influential guidelines about protocols and analysis strategies for such studies.
    a. Friedman, L., Stern, H., Brown, G. G., Mathalon, D., Turner, J., Glover, G. H., Gollub, R. L., Lauriello, J., Lim, K.O., Cannon, T., Greve, D. N., Bockholt, H. J., Belger, A., Mueller, B., Doty, M. H., He, J., Wells, W., Smyth, P., Pieper, S., Kim, S., Kubicki, M., Vangel, M., and Potkin, S. G. Test-retest and between-site reliability in a multicenter fMRI study. *Human Brain Mapping*, 2008; 29:958-972. PMC3670112
    b. Kim, S., Smyth, P, and Stern, H. (2010). A Bayesian mixture approach to modeling spatial activation patterns in multi-site fMRI data. *IEEE Transactions on Medical Imaging*, 2010; 29:1260-1274. PMC3690175
    c. Brown, G.G., Mathalon, D.H., Stern, H., Ford, J., Mueller, B., Greve, D.N., McCarthy, G., Voyvodic, J., Glover, G., Diaz, M., Yetter, E., Ozyurt, I.B., Jorgensen, K.W., Wible, C.G., Turner, J.A., Thompson, W.K., Potkin, S.G., FBIRN. Multisite reliability of cognitive BOLD data. *Neuroimage*, 2011; 54:2163-2175. PMC3009557
    d. Zhou, B, Konstorum, A, Duong, T, Tieu, KH, Wells, WM, Brown, GG, Stern, HS, Shahbaba, B (2013). A hierarchical modeling approach to data analysis and study design in a multi-site experimental fMRI study. *Psychometrika*, 2013;78:260-278. PMC4142354

4. **Statistical methods for missing data** – Unanticipated missing data values have a significant impact on health-related research. I have worked with statistics graduate students to develop novel approaches to performing statistical analyses in the presence of such missing data (publications a. and b. below). In addition, I served in 2009 and 2010 on a National Research Council (part of the National Academies) on approaches for missing data in clinical trials. The panel produced an influential report and the two summary articles identified below (publications c. and d. below).
    a. Sinharay, S., Stern H. S., and Russell, D. (2001). The use of multiple imputation for the analysis of missing data. *Psychological Methods*, 6, 317-329.
    b. Stern, H. S., and Jeon, Y. (2004). Applying structural equation models with incomplete data. In *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, eds. A. Gelman and X-L Meng. John Wiley and Sons: Chichester, UK, 331-342.
    c. Little, RJ, D'Agostino, R, Cohen, ML, Dickersin, K, Emerson, SS, Farrar, JT, Frangakis, C, Hogan, JW, Molenberghs, G, Murphy, SA, Neaton, JD, Rotnitzky, A, Scharfstein, D, Shih, WJ, Siegel, JP, Stern H. The prevention and treatment of missing data in clinical trials. *New England Journal of Medicine*, 2012; 367:1355-60. PMC3771340.

d. Little, RJ, Cohen, ML, Dickersin, K, Emerson SS, Farrar, JT, Neaton, JD, Shih, W, Siegel, JP, Stern, H (2012). The design and conduct of clinical trials to limit missing data. *Statistics in Medicine*, 2012; 31:3433-43. PMC5944851

5. **Bioinformatics** – I have worked on several projects involving the analysis of genomic data. One of these included developing novel methods for a sequencing approach that generated repeated short reads of genomic sequences. The statistical models developed by myself and a graduate student leverage the similarity of this problem to a familiar problem in ecology known as the species problem. A second problem concerned an approach to identify the nearest relative of a given organism in a known finite population of candidates. I have also worked more recently on genome-wide association studies.
    a. Zhang, H. and Stern, H. (2005). Investigation of a generalized multinomial model for species data. *Journal of Statistical Computation and Simulation*, 75(5), 347-362.
    b. Zhang, H, and Stern, H. S. (2006). Assessment of ancestry probabilities in the presence of genotype errors. *Theoretical and Applied Genetics*, 112(3), 472-482. PMID:16307226
    c. Zhang, H. and Stern, H. (2008). Inferences for genotyping error rate in ancestry identification from SSR marker profiles. *Journal of Agricultural, Biological and Environmental Statistics*, 14, 170-187.
    d. Zhang, H. and Stern, H. (2009). Sample size calculation for finding unseen species. *Bayesian Analysis*, 4(4), 763-792.

Complete List of Published Work in MyBibliography:
https://www.ncbi.nlm.nih.gov/sites/myncbi/1z7j7rJPkmhkO/bibliography/47999712/public/?sort=date&direction=ascending

D. Research Support

Ongoing Research Support

NIH – P50 MH096889 Baram (PI) 05/01/2013-04/30/19
Title: Conte Center on Fragmented Early Life Environment and Cognitive and Emotional Vulnerabilities
Goals: The Center will explore the role of environmental influences on development of cognitive and emotional disorders. The projects integrate longitudinal studies of a large human cohort with mechanistic animal studies. The central hypothesis is that fragmented and unpredictable early-life experiences increases the vulnerability of developing individuals to cognitive and emotional disorders. Brain imaging and behavioral studies are used to characterize fragmented behavior and its consequences.
Role: Senior Investigator (Director of Biostatistics, Computing and Data Management Core)

NIST – Cooperative Agreement 70NANB15H176 Carriquiry (PI) 06/01/2015-05/31/2020
Title: Center for Statistical Applications in Forensic Evidence (NIST Center of Excellence)
Goals: The Center investigates the scientific underpinning of the evaluation of forensic evidence with a focus on pattern evidence (latent prints, shoe prints, etc.) and digital evidence. Individual projects attempt to identify probabilistic and statistical models that can be used to quantify the strength of the evidence and also explore the ways in which the evidence interpretation can be explained to participants in the criminal justice process (attorneys, judges, jurors).
Role: Principal Investigator for UC Irvine subcontract

California Initiative to Advance Precision Medicine (CIAPM) Greenfield (PI) 11/01/2016-5/30/2018
Title: Precision Medicine for Early Prostate Cancer: Integrating Biological and Patient Complexity Variables to Predict Treatment Response
Goals: Project will develop approaches for personalizing prostate cancer treatment by incorporating information regarding: (a) patient characteristics and patient reported outcomes such as socio-demographic information, health status and disease management burden, (b) traditional prostate cancer severity indicators and (c) an existing genomic test that estimates the probability of cancer spread after surgery.
Role: Senior Investigator