Conversations with Herman Rubin

Mary Ellen Bock[1]

Purdue University

Abstract: Herman Rubin was born October 27, 1926 in Chicago, Illinois. He obtained his Ph.D. in Mathematics from the University of Chicago in 1948 at the age of 21. He has been on the faculty of Stanford University, the University of Oregon, Michigan State University and Purdue University, where he is currently Professor of Statistics and Professor of Mathematics. He is a Fellow of the Institute of Mathematical Statistics and of the American Association for the Advancement of Science as well as a member of Sigma Xi.

He is well known for his broad ranging mathematical research interests and for fundamental contributions in Bayesian decision theory, in set theory, in estimations for simultaneous equations, in probability and in asymptotic statistics.

These conversations took place during the 2003–2004 academic year at Purdue University.

Herman, it is great that the IMS is bringing out this Festschrift for you. I am delighted to be able to prepare this interview with you. I guess we always want to know about childhood. So Herman, where did you grow up?

I was born in Chicago, Illinois, and grew up there, the oldest of three children. Both of my parents were immigrants, my father from Russia and my mother from Russian-occupied Poland. My mother's sister was also an immigrant and she taught me to read at the age of three.

What was your educational background? Did you receive special training in mathematics?

I went to the Chicago public schools for grammar school and was a voracious reader in the public library. But the material was organized by grade level and I did not find much on mathematics beyond arithmetic. But the summer before I went to high school I discovered algebra when I came upon a book about it while visiting New York City. After reading the book, I tested out of algebra in the first month of the first year of high school. In high school I found many more advanced books about mathematics in the public library; I taught myself material through calculus while taking plane geometry in high school. After two years at the public high schools, I was given a scholarship to a combined high school/college program at the University of Chicago. I could have graduated high school after a total of three years but delayed the official high school graduation by one year because I could take more college courses and not pay the college tuition. I received the high school diploma in June of 1943, the bachelor degree SB (Mathematics major with Physics minor) in December of 1944 and the master degree SM (Mathematics) in March of 1945. At the University of Chicago, almost all of my courses beyond the bachelor's were in abstract mathematics but my Ph.D. dissertation was in statistics.

How did you get interested in the field of statistics if most of your courses were in abstract mathematics?

My interest developed during my stint at the Cowles Commission for Research in Economics (CCRE) which was housed at the University of Chicago. In 1944

---

[1]Department of Statistics, Purdue University, 150 N. University Street, West Lafayette, IN 42907-20689, USA. e-mail: `mbock@purdue.edu`

CCRE needed a mathematics research assistant because their current assistant was being drafted into the U.S. military. At the time I was a student in the undergraduate/graduate program at the University of Chicago and, aside from my mathematical abilities, one of my qualifications was that I was too young to be drafted. So in July of 1944 at the age of seventeen I became a research assistant for CCRE.

I became interested in statistics because the leader of CCRE, Jacob Marschak, who took over in 1943, had decided to concentrate the work of the group on the problems of stochastic simultaneous equations found in economics.

Who worked with you when you joined CCRE?

My initial work was with Tjalling Koopmans who had joined CCRE at the same time as I. He was brought in to concentrate on the mathematical aspects. My first paper was a solution to a problem of Koopmans for the approximate distribution of the circular serial correlation coefficients under the null hypothesis and it appeared in the *Annals of Mathematical Statistics* in 1945. The main problem I worked on with Koopmans was to estimate the parameters of a system of stochastic equations including lags and to derive their properties. (Individual equations might have more than one dependent variable and least squares was already known to be inconsistent when applied individually to each equation.) I developed some Maximum Likelihood techniques and their properties for the time series lags to attack the problem.

I understand that the work at CCRE was interrupted.

The work was interrupted because I was drafted into the U.S. Army in March, 1945, at the age of 18. The bulk of the work I mentioned was published as a joint paper by myself with Koopmans and Roy Leipnik. (Roy was a research assistant in CCRE from February, 1945, to July, 1946, and took over the work with Koopmans after I was drafted.)

I was discharged from the Army in December, 1945, and returned to the University of Chicago as a graduate student and CCRE as a research assistant in January, 1946. (CCRE promoted me to research associate in November, 1946.)

Who worked with you on your return to CCRE?

I began to work with Theodore W. Anderson who had joined the CCRE as a research associate in November, 1945, in my absence. One source of inspiration for our work was a talk I heard after my return given by the biologist Sewall Wright. (He had given a general formulation for the problem of solving simultaneous stochastic equations in 1919.) I realized that factor analysis was another example of simultaneous stochastic equations and this led to a paper on it with Anderson.

Anderson and I collaborated on three papers. The first paper developed the maximum likelihood estimator of the coefficients of a single equation in a system of stochastic equations; the estimator is now known as the Limited Information Maximum Likelihood (LIML) estimator. The second paper developed the large-sample distribution theory. The LIML estimator had been developed in Anderson's 1945 dissertation. Our third joint paper developed maximum likelihood methods for factor analysis models with different identification conditions. It was a pretty innovative paper at the time.

Another source of interesting questions was Meyer A. Girshick. Early in 1946 Koopmans gave me a letter from Girshick about the problem of estimating a single equation (with more than one dependent variable) without estimating the entire complete system of equations. (A system of equations is complete if there are enough equations of the right sort so that all the coefficients could be consistently

estimated, essentially a multivariate regression problem.) I developed it somewhat and then collaborated on further aspects with T. W. Anderson. This work (with credit to Girshick) appeared finally in 1949 and 1950 in the *Annals of Mathematical Statistics.* The publication was somewhat delayed because in those days it was a major job (without the benefit of email) to communicate with the referees and my coauthor Anderson who was in Sweden during the 1947–1948 academic year. (He left CCRE in September, 1946, to go to Columbia.)

What about the Ph.D. degree?

I received the Ph.D. degree from the University of Chicago in March, 1948, at the age of 21. My official advisor was Paul Halmos in the Department of Mathematics. The dissertation topic grew out of my work at CCRE. It involved extending the original problem of Girshick of estimating a single equation to that of estimating a subsystem of equations without estimating the entire complete system of equations. The dissertation was typed up while I was on leave from CCRE as a post-doc at the Institute for Advanced Study in Princeton during the academic year of 1947–1948.

You have made major contributions to the field of asymptotics. Why do you feel asymptotics are important?

The need for asymptotics at CCRE inspired me and this culminated in my first major insights in 1949. Some of my contributions to the problem were the asymptotic theorems on limiting distributions which were never published. I introduced the idea of a random function into the generalization of the Slutsky Theorems. James Hannan and Vaclav Fabian gave the proofs in their book crediting me. For inspiration I used general topology (although metric topology is adequate). For me, the more I generalize the problem to an area of abstract mathematics the easier it is for me to understand it since I can get rid of the part which doesn't add to the meaning of the problem. Even when I computed something, if I could generalize it, then it led to insight. I know that is not how most people like to do mathematics.

You have had an abiding interest in computing. What was it like then?

At CCRE, computations were done with electromechanical desk calculators and a staff of three operated the calculators. Computations BC (Before Computers) were much slower. I was in charge of computing at CCRE until Herman Chernoff took over when I left for Princeton in August of 1947; we had some pretty funny experiences making the equipment work. (He had come there as a research assistant to CCRE in July, 1947.)

You are well known for your interests in statistical decision theory. Was it influenced by the CCRE experience?

The CCRE emphasis on economics was a factor. The idea of a utility scale for actions assuming that the state of nature is fully known, which goes back much farther, was important in quantitative economics for a long time. However, no essential progress had been made in getting a clear scale until the von Neumann – Morgenstern axioms for cardinal utility appeared in their book "Theory of Games and Economic Behavior" in 1944. One of their key contributions was the use of randomization. Researchers at the CCRE in 1947 were considering extending the ideas to unknown states of nature while I was there. I observed that adding one simple axiom made the utility for unknown states of nature a positive linear functional of the utility functions indexed by the given states of nature. (This is essentially the prior Bayes approach.)

In the early years of decision theory, the main progress was made in proving theorems and refining the concepts, and I had my share in this. Stanford was a

center of activity in this and I went there after leaving CCRE. I had various degrees of collaboration with Blackwell, Girshick, Karlin, and Chernoff, and numerous discussions with Stein. Four dissertations on decision theory were written under me.

You have worked on a variety of problems in probability, particularly stochastic integration, characterizations and infinite divisibility. You have collaborated with numerous people on these. How was that experience?

Yes, I have collaborated with C. R. Rao, Burgess Davis, Tom Sellke, Anirban DasGupta, Steve Samuels, Prem Puri, Rick Vitale and many others on questions in probability. The results with C. R. Rao got to be known as Rao-Rubin theorems; we were both visiting Stanford that year. Burgess Davis and Jeesen Chen asked an interesting question about uniform empirical processes. Tom Sellke and I worked on several Choquet type decomposition problems in the eighties. I have always enjoyed using characteristic functions as a tool, as those works did. I am glad my book length review with Arup Bose and Anirban DasGupta on infinite divisibility got published a couple of years ago; we worked many years on that one. With Prem Puri and Steve Samuels, the works were more in applied probability, but they were good problems. And, you mention stochastic integration. Yes, I too had thought of the Stratonovich integral. I gave a talk introducing the idea behind the Stratonovich integral at the IMS meeting in Seattle in 1956. My Ph.D. student Don Fisk later wrote a thesis on it in 1961. I myself did not write it up or pursue it formally. Probability questions are always interesting.

In the fifties, you collaborated with Karlin on introducing monotone likelihood ratio. This has had a very major impact. How did that idea originate?

Steve Allen, a Ph.D. student under Girshick, had come up with a proof that in the exponential family, monotone procedures are essentially complete for one-sided testing problems. I first wrote a technical report. Karlin and I realized it works for monotone likelihood ratio. We generalized that result of Allen and gave applications. Yes, it later led to concepts such as total positivity. Karlin has written much about it.

What did you do with Chernoff?

That was the beginning of my interest in the discontinuous density problems; we had a paper together in the third Berkeley symposium. But the relationship extended beyond professional collaboration.

You have a number of publications in set theory. How did this interest arise?

I was always interested in set theory and while in graduate school at Chicago I took a course from the topologist John L. Kelley which piqued it even more. There is a version of set theory that he showed me (called the Morse-Kelley set theory) which is stronger than the usual set theory because you can prove the consistency of the usual set theories (such as the Zermelo-Frankel or the von Neuman-Bernays-Godel) in the Morse-Kelley system.

From CCRE I went to Stanford's Department of Statistics in 1949 as an Assistant Professor and eventually met Jean Hirsh when she arrived later as a mathematics Ph.D. graduate student in logic there. We married in 1952. Her interests in logic and mine in set theory eventually led to a professional collaboration.

Later Pat Suppes was teaching a class on set theory for which I gave some lectures on the axiom of choice. Professor Suppes who knew both of us suggested

that Jean and I write a book on the various equivalents of the axiom of choice. (Jean received the Ph.D. in mathematics for her work in logic in 1955 and Suppes was her advisor.) After at least eight years, two moves and two children, we finally finished the book.

With two parents with Ph.D.'s in mathematics, were the children also interested in mathematics?

Arthur was the oldest (born in 1956) and went on to get a Ph.D. in mathematics from California Institute of Technology (at the age of 22) after being a Putnam Fellow four times. Arthur and Paul Erdós wrote a paper together. Leonore who was born in 1958 received a bachelor's degree with honors jointly in mathematics and chemistry from Michigan State University and went on to get a Ph.D. in chemistry from Carnegie Mellon.

You mentioned several moves. Where did you go?

After I left Stanford in 1955, I went to the Department of Mathematics at the University of Oregon for four years. (Because of nepotism rules, Jean was not allowed to have a regular position at the same university or even paid by the State of Oregon.) I had some collaborations with Howard Tucker and A.T. Bharucha-Reid at Oregon. From Oregon I went to Michigan State University's Department of Statistics in 1959. Again, Jean could not be hired because of nepotism rules. The set theory book on the axiom of choice by Jean and me was published while we were at Michigan State.

Most of my collaborations at Michigan State were with Martin Fox in decision theory, game theory and functions of Markov states. It was also the start of a collaboration with J. Sethuraman. We did some work on what is now called moderate deviations. We also later collaborated on Bayes risk efficiency.

Then in 1967 we both came to Purdue where Jean received an offer from the Department of Mathematics that included tenure. I joined the Department of Statistics and the Department of Mathematics as a full Professor and Jean joined Math as an assistant professor. I have been here ever since and my wife Jean was a full Professor of Mathematics here until her death in 2002. She is honored by an annual seminar and remembered for her support for women faculty in academia. She started a scholarship fund for mathematics students in her will.

One of your strong ongoing interests is in prior Bayesian robustness. How do you describe it?

One of the difficulties of Bayesian analysis is coming up with a good prior and loss function. (I have been saying for years that the prior and the loss cannot be separated. The Carnegie Mellon school is doing some work on that now.) When I talk about prior Bayesian robustness I assume that one does not yet see the random observation X whose distribution depends on the unknown state of nature. One considers the choice of different priors for which one averages over the possible states of nature and over the possible random observations. This is different from posterior Bayesian robustness in which one considers the choice of different priors given the random observation X whose distribution depends on the unknown state of nature. If you can get posterior Bayesian robustness, then you automatically get prior Bayesian robustness but seldom are we so lucky as to find posterior Bayesian robustness. It is actually the axioms of utility that decree we should worry about prior Bayesian robustness. When I am faced with a choice among priors, all of which seem about the same to me, then I am very concerned about the possible alternative consequences of applying either one if it is drastically wrong. For instance, suppose I

am using squared error loss to estimate the mean of a normal random variable with variance one. The first prior for the unknown mean might be normal with mean zero and standard deviation 10 while a second prior for the unknown mean might be normal with mean zero and standard deviation 1000. Now using the first prior could be disastrous (in terms of a loss that is averaged over the values of the state of nature as well as the possible mean values) when the second prior is appropriate. Yet using the second prior would not be so bad if the first prior were appropriate. In contrast for the posterior Bayesian robustness approach, if the observation X is large, then the posterior loss is bad in either case if the wrong prior is used.

You have had a long term interest in random number generation. What inspired you?

I heard that a professor at Columbia had announced to his class that he would give a midterm in each of five three-week periods, the particular week to be chosen at random by tossing a coin. Finding an efficient way to do this was an interesting problem to me. I observed that generating all five results at once was far more efficient that generating the results one at a time from a discrete distribution. This eventually led to less trivial questions and was the start of my interest in efficient methods for generating random numbers.

The two main problems I find interesting are the following: how to get lots of random numbers which are independent and uniform; and how to turn them into independent random numbers from some other distributions.

In the case of the first main problem, when generating independent uniform random variables, most people use pseudo random numbers. It is almost impossible to prove that they have all the desired properties. Of course, they fail the test that they come from the pseudo random number generator! For physical random numbers, one can question the accuracy of the model for the physical process. (I have a technical report about paradoxes caused by the effect of dead time.) My personal preference is to use a stream of physical random numbers and a stream of pseudo random numbers to produce a stream of random numbers whose qualities should be at least as good as either of the original two streams.

In the case of the second main problem, even when you have independent uniform random variables, the problem of using them to generate variables from other distributions is sometimes hard. The basic issue of efficiency is not the question of the number of bits used but rather the computational cost... and this is a complex question. I have some technical reports on these issues.

Computing issues in probability and statistics have been a topic of research for you, too. What are your comments?

Computation is an obvious issue in the generation of random numbers. But computation of probabilities is also important and it is often best done through the use of characteristic functions (i.e. Fourier transforms). I find that reasonably efficient computational procedures require complex integration... and that requires a knowledge of analytic functions.

Another important area is the computation of Maximum Likelihood estimates (MLE). This typically requires more expertise in analysis than is usually expected in most statistics graduate programs. For Bayes procedures, integration computing problems are commonplace. Many have pointed out the difficulties in posterior Bayes computations. Simulation is another area of computing problems. In 1970, I was using simulation to compute theoretical expectations for Kolmogorov-Smirnov and Kuiper statistics under nonnull hypotheses. The finite sample distribution was approximated by a modification of a Brownian Bridge. My first observation was

that I could not just simulate at a finite number of points because the Brownian Bridge changed too rapidly. This was handled by simulating the max and the min in various intervals independently (even though the max and min are not independent) and it worked quite well. The reason it worked well was that the probability that the max and min of the whole process are both in the same interval is extremely small. This points out that it is often necessary to do analysis before numerical analysis.

How has the internet affected your work?

I contribute to newsgroups and give advice to those that ask. I also join extended discussions on how things should be done. In general, I find out things that I might not otherwise know because the internet puts me in contact with lots of people. It is much easier to collaborate and I think it is a great advantage for research. You see, you can put a hard problem on the web and get help from experts.

I have not yet directly engaged in internet communication in producing papers, but I have been a joint author with local collaborators who have. It would help if we could have an easier way to communicate mathematics notation.

What are you working on these days?

Hui Xu, a Ph.D. student, is doing some work with me on density estimation when there are discontinuities. I have had a longstanding interest in that. I am also doing some work on random number generation with Brad Johnson, another Ph.D. student. And I just finished a paper on the Binomial n problem with Anirban DasGupta; it is coming out early next year in the Chernoff Festschrift!

What else do you enjoy? Have you done much traveling?

Well, I enjoy going to the concerts and the operas, although I did not so much as a child. I try to keep track of what is going on in mathematics. Previously, I could only do it by picking up the journals at the library. Now you can do some of it by using the net. I think what I enjoy the most is talking to students and people and be of any help that I can. You see, I have an open door policy.

For me, the most rewarding part of traveling is talking to people about interesting questions. I enjoyed going to the International Congress of Mathematicians at Stockholm, the Oberwolfach meeting I went to, and a meeting at Israel. I went to the ISI in 1974. Mahalanobis had just passed away, but there were a lot of people from everywhere at that meeting. I remember Persi Diaconis being there and Peter Bickel and many others. Urbanik was eating the raw jalapenos for his snacks. But even the symposium food was too spicy for me. From Calcutta, I went to Delhi. B. K. Kale invited me to come to Jaipur. It was an interesting trip. I went by a private limousine and returned on a public bus! Anirban wants to take me back there. We will see.

Have your many years of teaching influenced your ideas about statistical education?

Definitely. I believe that it is the unusual person who can go easily from the specific to the abstract. I think it is easier to go from the abstract to the specific. (Most of my colleagues disagree with me on this.) I have no objection to using examples after a concept. But going from special cases to the general still leaves the need for unlearning, which is difficult.

Because theorems and proofs are an important part of mathematical statistics, I believe that students who did not have some kind of course with theorems and proofs in high school, say Euclid-type geometry, flounder when they reach mathematics in college. We must improve quality of mathematics education in the US.

Competition is getting very strong and economic health is directly related to quality of education.

But even more important than experience with theorems and proofs are courses that emphasize concepts. For instance, thinking about integration as a limit of a sum is a crucial idea in statistics, especially for expectations. Students have difficulty if they learn integration as antidifferentiation, i.e. the "opposite of differentiation," and not as a summing process. I believe that it is possible for students to learn concepts directly if properly explained. This does not mean that a student will be able to use the concept upon hearing the words. Considerable learning may need to occur before the "light bulb" goes on.

To close the interview, what do you have to say about the future of statistics?

The biggest opportunities lie in the development of decision theoretic approaches to the problems of individual users where one considers ALL the consequences of the proposed solutions. Taking all the consequences into consideration can produce very difficult mathematical problems and provides great opportunities for those with mathematical expertise.

This is in contrast to the emphasis today on the development of general recipes that are used for solving problems and that are often used inappropriately. The latter two-thirds of the nineteenth century saw a similar emphasis. The turnaround came after World War II with people going into statistics from good mathematics programs who could attack the challenging mathematical problems. Before the turnaround there was also a rush by users, as now, to use statistical methods without understanding the assumptions and their consequences. I feel it is the user who must make the assumptions rather than just the statistician! Arguably, in a quantitative area, the user is not well prepared to do that.

Both those becoming statisticians and the users need to realize that there are underlying concepts for the field and they must use an understanding of the concepts rather than a catalog of methods. Just knowing how to compute does not help, and even being able only to prove lots of theorems would not. We CAN teach these concepts, and many of them even at fairly low level courses. The applied statistician needs to be able in many cases to invent new methods on the spot. There will be great opportunities for collaboration between applied scientists and mathematicians in the coming years. I hope neither ignores the other as an ancillary. That will be a mistake.

Thank you for your interesting views on research and teaching and for the interesting stories on your life. It was a pleasure. Good luck to you and we hope to continue to walk through the door and ask a question and get help. You have been a gracious resource to all of us. I wish you health and happiness.

You are very welcome.

References

[1] "On the distribution of the serial correlation coefficient," *Annals of Mathematical Statistics*, Vol. **16**, 1945, pp. 211–215.

[2] "The theory of decision procedures for distributions with monotone likelihood ratio," *Annals of Mathematical Statistics*, Vol. **27**, 1956, pp. 272–300. With S. Karlin.

[3] "Distributions possessing a monotone likelihood ratio," *Journal of the American Statistical Association*, Vol. **51**, 1956, pp. 637–643. With S. Karlin.

[4] "Measuring the equation systems of dynamic economics," *Cowles Commission Monograph*, No. **10**, 1950, pp. 53–237. With T. C. Koopmans and R. B., Leipnik.

[5] "Consistency of maximum-likelihood estimates in the explosive case," *Cowles Commission Monograph,* No. **10**, 1950, pp. 356–364.

[6] "Admissibility of quantile estimates of a single location parameter," *Annals of Mathematical Statistics*, Vol. **35**, 1964, pp. 1019–1030. With Martin Fox.

[7] *Equivalents of the Axiom of Choice II*, 1985, North Holland. With Jean E. Rubin.

[8] "How non-uniform can a uniform sample be?" *Probability Theory and Related Fields*, **73** (1986), 245–254. With Jeesen Chen and Burgess Davis.

[9] "On the distributions of sums of symmetric random variables and vectors." With Thomas Sellke. *Annals of Probability*, **14** (1986), 247–259.

[10] [10 Bose,A., DasGupta,A. and Rubin,H. (2002). A Contemporary Review of Infinitely Divisible Distributions and Processes, *Sankhȳa, Ser. A.* **64**, Part 3(ii), 763–819.

[11] "Estimation of the parameters of a single equation in a complete system of stochastic equation," *Annals of Mathematical Statistics*, Vol. **20**, 1949, pp. 46–63. With T. W. Anderson.

[12] "The asymptotic properties of estimates of the parameters of a single equation in a complete system of stochastic equations," *Annals of Mathematical Statistics*, Vol. **21**, 1950, pp. 370–382. With T.W. Anderson.

[13] "Statistical inference in factor analysis," *Proc. of the Third Berkeley Symp. on Math'l. Stat. and Prob'y.*, Vol. **5**, 1956, pp. 11–150. With T. W. Anderson.

[14] "Uniform convergence of random functions with applications to statistics," *Annals of Mathematical Statistics*, Vol. **27**, 1956, pp. 201–204.

[15] "Construction of random variables with specified distributions," *MSU Research Memorandum*, No. **88**, November 1961.

[16] "Some fast methods of generalizing random variables with preassigned distributions. I. General Acceptance–Rejection Procedures." Technical Report #466, Purdue University, Department of Statistics.

[17] "The effect of dead time on the physical generation of random digits." Technical Report #467, Purdue University, Department of Statistics.

[18] An Efficient Method of Generating Infinite-Precision Exponential Random Variables. Technical Report #86-39, Purdue University, Department of Statistics.

[19] Generating Non-Uniform Random Variables: Infinite Precision Procedures and Computational Complexity. Technical Report #86-58, Purdue University, Department of Statistics.

[20] "Probabilities of moderate deviations," *Sankhya A*, Vol. **27**, 1965, pp. 325–346. With J. Sethuraman.

[21] "Bayes risk efficiency," *Sankhya A*, Vol. **27**, 1965, pp. 347–356. With J. Sethuraman.

[22] "On a characterization of the Poisson distribution," *Sankhȳa A.* Vol. **26**, 1964, pp. 295–298. With C. R. Rao.

[23] Chen Z., Rubin H. and Vitale R. A. (1997). Independence and determination of probabilities. *Proc. Amer. Math. Soc.* **125**, No. 12, 3721–3723.

[24] "A characterization based on the absolute difference of two i.i.d. random variables," *Ann. Math. Stat.* **41** (1970) 2113–2122. With P. S. Puri.

[25] "On a characterization of the family of distributions with constant multivariate failure rates," *Annals of Probability.* **2** (1974) pp. 738–740. With P. S. Puri.

[26] "The Finite Memory Secretary Problem," Purdue University. With S. M. Samuels. *Annals of Prob.* **5**, 4 (1977) 627–635.

[27] Casella, G., DasGupta, A., Delampady, M., Rubin, H., and Strawderman, W. (2001). A new approach to default priors. *The Canadian Journal of Statistics.* **29**, No. 3, 437–450.

[28] DasGupta, A., and Rubin, H. (In Press). Estimation of the Binomial parameters when both parameters are unknown. *Special Issue of Journal of Statistical Planning and Inference in Honor of Herman Chernoff.*

[29] "Estimating the parameters of a differential process", *Annals of Mathematical Statistics*, Vol. **30**, 1959, pp. 641–658. With H. G. Tucker.

[30] "Generating functions and the semigroup theory of branching Markov process," *Proceedings of National Academy of Science*, Vol. **44**, 1958, pp. 1057–1060. With A. T. Bharucha-Reid.

[31] "The estimation of the location of a discontinuity in density," *Proceedings of the Third Berkeley Symp. of Math'l. Stat. & Prob'y.*, Vol. **1**, 1956, pp. 19–37. With H. Chernoff.

[32] Rubin, H. (1987). A weak system of axioms for "rational" behavior and the non–separability of utility from prior. *Statistics and Decisions* pp. 47–58.