Statistical Science
2022, Vol. 37, No. 3, 443–454
https://doi.org/10.1214/21-STS845

A Conversation with Stephen Portnoy

Xuming He and Xiaofeng Shao

Abstract. Steve Portnoy was born in Kankakee, Illinois in 1942. He did his undergraduate studies in mathematics at Massachusetts Institute of Technology, and then earned a master's degree and a Ph.D. degree from the statistics department at Stanford University in 1966 and 1969, respectively.

Steve Portnoy has had a distinguished career and is widely recognized as a preeminent mathematical statistician. He has made pioneering and influential contributions in several areas in statistics, including asymptotic theory, robust statistics, quantile regression, and statistics in biology. He has published more than 100 research articles. He is a former co-editor of *Journal of the American Statistical Association (Theory and Methods)*, an elected fellow of American Statistical Association (ASA), Institute of Mathematical Statistics (IMS) and American Association for the Advancement of Science (AAAS).

Steve's professional positions have included being on the faculty of the Department of Statistics at Harvard University and the University of Illinois at Urbana-Champaign for more than 30 years. He was a founding member of the Department of Statistics at the University of Illinois in 1985 and served as the division chair for the Statistics Program in the Mathematics department before the Statistics department was established.

Key words and phrases: Asymptotic approximation, high-dimensional models, quantile regression, robust statistics.

The following conversation took place prior to the Workshop "Asymptotic Theory, Robust Statistics, and Quantile Regression: A Workshop Celebrating the Contributions of Stephen Portnoy" that took place at University of Illinois at Urbana-Champaign (UIUC) on October 16, 2021. The interviewers are abbreviated as (**XH**) and (**XS**) for Xuming He and Xiaofeng Shao, respectively. The interview has been edited for clarity but may contain informal phrases and spontaneous responses at times.

1. STUDENT DAYS

XH: Can you tell us your experience at MIT as an undergraduate, and what led you to study statistics at Stanford?

SP: When I entered MIT (Figure 1), I was well prepared in math and took second year calculus as a freshman. I took a rigorous undergrad real variables course in the next year and the graduate measure and real variables course as a junior (along with differential geometry, point-set topology, and an early AI course). As you can see, MIT encouraged students to go as fast as they wanted. Unfortunately, I was going too fast. I realized I needed to complete my math requirements at more of an undergrad level. Also the AI and Econ courses I was taking piqued my interest in applied mathematics, which led me to take an excellent Introduction to Mathematical Statistics course from Harold Freeman. The course was sufficiently great to convince me to change from graduate work in pure math to graduate work in statistics. I am quite sure that Freeman's recommendation was crucial for my being accepted by Stanford.

MIT was really great—the faculty was remarkably accessible (even for undergrads); and most important, I met Esther as a senior at the MIT Science Fiction Society.

XS: Did you recall any lectures that you particularly enjoyed at Stanford?

SP: The courses I remember most were Charles Stein's. Charles taught a decision theory course that enthralled me. He was a very systematic lecturer—though rather abstract even at Stanford. For any theorem, he tended to present an approach that best clarified why the result held. In the second year, he was teaching decision theory again and asked me to write up notes (I still have them). This really cemented my knowledge of the basic results—especially admissibility (basically advanced convex analysis) and invariance (using the theory of transformation

Xuming He is H.C. Carver Collegiate Professor, Department of Statistics, University of Michigan, Ann Arbor, USA (e-mail: xmhe@umich.edu). Xiaofeng Shao is Professor, Department of Statistics, University of Illinois at Urbana-Champaign, Champaign, USA (e-mail: xshao@illinois.edu).

FIG. 1. *Steve at MIT dorm around* 1960.

groups). Clearly, Charles preferred a rather abstract approach to the subject—but to me this was great. As an aside, Charles taught the first semester of ANOVA at the same time using the coordinate-free approach—basically it was a course in finite dimensional Hilbert Spaces! Our introduction to balanced incomplete block designs was basically the theory of two noncommutative linear operators (which we eventually learned were the projections providing the block averages and the within block contrasts). Fortunately for students who wanted to learn how to apply ANOVA, the second semester was taught by Lincoln Moses, where we covered things like split-plot designs, random-effect and mixed models, etc.

There were also memorable courses by Ingram Olkin on Multivariate Analysis, and courses on Probability and Applied Probability by Rupert Miller. Chernoff taught large sample theory and sequential methods, and there were multiple courses on Time Series at Stanford. Overall, the quality of course work was outstanding. I still have many of my class notes, and still use them in my research.

XS: Anyone from your cohort whom you would like to tell us about?

SP: Of course, one generally learns a great deal from classmates. Carl Morris and Jim Zidek were finishing their theses under Charles when I was first starting—and both were extremely helpful. Certainly, the quality of students (as well as faculty) generated excitement and motivation.

XS: What was your thesis research at Stanford? What was like working with Charles Stein?

SP: We met weekly. Charles was exceptionally innovative, and often brought up unique novel approaches to various results. But my thesis research was mainly self-motivated. I was working on extending a sufficient (and nearly necessary) condition for admissibility that Charles developed for one-dimensional parameters. The case of multiple parametric dimension was open. Basically, I would tell him what I was thinking about, and every so often he had a crucial (generally quite brief) comment that was extremely helpful. But mainly, I was left to work on my own, which was common at Stanford; they expected all grad students to progress with only a few nudges now and then. This meant that although the thesis wouldn't have been possible without Charles, almost all my work was original. When I finally had a theorem, I applied it to estimating random effects, and that was my thesis: one theorem and one example.

XH: Was statistics more closely related to mathematics back then? Today do you think a Ph.D. student in statistics needs to come from mathematics majors?

SP: Certainly back then I think it was much more common for academic statisticians to have very strong math backgrounds. Not all would have been math majors, but most of them had somewhat more than calculus and linear algebra. And I think that's still necessary today. I don't think you can really understand the statistical theory that Ph.D. students need without having at least some introduction to measure theory and probability.

XH: To follow up on what you said that even back then not everybody came from math, what were the academic background of those students? Did anyone come from physics, for example? I was thinking about Bob Wijsman who was a professor at University of Illinois at Urbana-Champaign (UIUC).

SP: I don't remember any graduate students who started in physics. Mostly out of social and biological sciences or medicine. Those students working in other areas who decided to do graduate work in statistics recognized that the department of statistics required a good bit of math. They either took it before they got here or expected to take a good bit of math when they got here. Even at Stanford there were some people in the biostatistics seminar, I don't remember their names, but I am pretty sure they didn't come from math. There were students who definitely felt the measure theory courses they were required to take at Stanford were fairly advanced for them. Some weren't really prepared for them, but those courses were required.

XH: I see. That's interesting because I thought that if you didn't come from math, you would not have any real analysis or measure theory background, and it would be very difficult to finish a Ph.D. in statistics.

SP: Yes, I think it was difficult without some advanced real variable courses at the graduate level. I think some students in other areas might pick it up when they realize they might need it. But otherwise, they would need the prerequisites in the first year, and this would delay taking the measure and probability courses. I think it was not uncommon for people like me to have a rather deep math background. Even at Stanford, I was probably more mathematical than most other students. I did take several more advanced courses in the math department. Karel DeLeeuw was a great teacher, and I took courses in topological groups and advanced functional analysis from him.

XH: Karel DeLeeuw?

SP: He was the professor who was murdered by a disgruntled former student. It was traumatic for people who knew him.

2. PERSONAL LIFE

XS: Where was your home town? Could you tell us a little bit about your family?

SP: I was born in Kankakee, Illinois, to parents both of whose families came from Eastern European Jewish immigrants. In fact my father was born in what is now the Ukraine—and so I am also a first generation American. I have a younger brother, who also had mathematical interests and got a MS in statistics at UCLA before going to work for the County of Los Angeles. My mother would have liked to go to college, but her family simply didn't have the resources during the depression, and she had to go to work. My father did attend UIUC (I think for only one semester) but finished his education at a 2-year community college in Kankakee. He became a successful business man; but I think he could have become a professor (his main interest was history) if resources were available. Both of my grandfathers had small clothing businesses that simply didn't really provide enough money for college, especially during the depression. My father's father was in fact a tailor in Kankakee—Portnoy means tailor in Russian.

XH: You mentioned earlier that you liked math even before college. Did you have a favorite hobby in high school?

SP: I did some coin collecting. Not really seriously. Mainly I did a lot of reading in science, playing with my chemistry set, my telescope, my microscope, just being an all around nerd. (all laugh)

Actually, I have always had a strong interest in music, mainly classical. I played some piano in my early teens, and Esther and I enjoyed concerts at the Krannert Center for Performing Arts: a truly great venue and one of the great attractions of Urbana-Champaign. Esther was actually a music major for a short while, in piano, oboe, and voice, and my daughter Rachel has a MA from UIUC in voice and choral conducting. About 10 years ago I decided to take violin lessons from a friend, Robin Kearton, the director of C4A, the Center for the Arts in Urbana (Figure 2). C4A offered group lessons for children, and some of the parents formed a band to have something to do while waiting for their children to finish. They called it the Cretaceous Band (the logo was a T Rex playing a trombone).

XH: What about hiking? I have known you as a hiker. When did you start serious hiking?

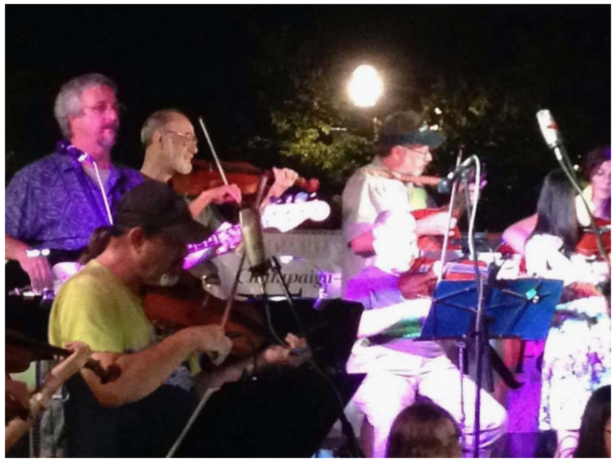

FIG. 2. *Steve playing violin in a gig in* 2015.

SP: I didn't really start that until I got to Boston. When I was an undergraduate, I walked around Boston a good deal.

Esther and I are both avid hikers. She was a student at Boston University and came with her roommate who was dating an MIT student who was active in the Science Fiction Society. After the meeting, most of us walked about 1/2 mile to a pizza place in Central Square. Esther and I sort of separated from the group and were walking together. About 58 years later, we still do lots of great hiking together.

XH: Tell us your most memorable hiking experience. Or your longest hike?

SP: There are many, Haha.... One of the longest happened just a couple of years ago; we went on a coast-to-coast hike in northern England, on the Wainwright trail. We didn't go all 192 miles. We had eight days and there was a group of six of us with a guide. We went about 10 miles a day.

XS: Wow.

SP: Yes, it was good bit of hiking. That was quite enjoyable. We would try to build in a hiking holiday whenever we flew to Europe. That was almost once a year. We would take time out there either before or after a conference, do some hiking. We hiked in the Alps several times (Figure 3) and in England and Scotland and the Czech Republic, Slovakia, Greece, and Germany. We spent two sabbaticals in Australia, where we did a great deal of hiking (I once estimated that if you added up the distance and elevation gain of all the hikes in Australia, it would be equal to hiking from the Indian Ocean to the top of Everest!).

XH: Yeah. That sounds impressive.

XS: When did your interests in mathematics and statistics begin to emerge? What were the most influential factors?

SP: I was always good at math and really enjoyed solving intellectual problems, especially puzzles. I don't re-

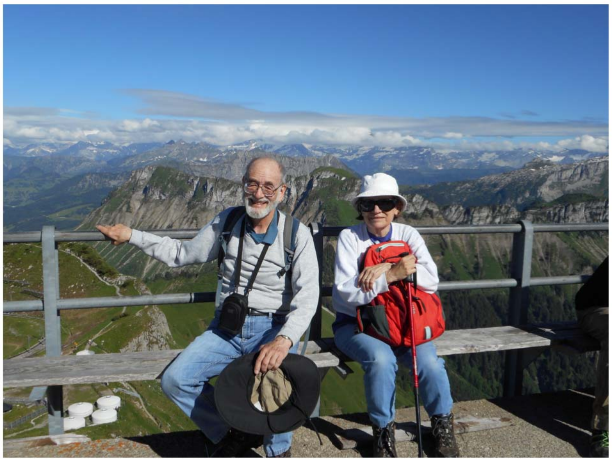

FIG. 3. *Esther and Steve Portnoy hiking in the Alps.*

member my parents even mentioning college, but by middle school I remember taking to a friend at summer camp who asked what I wanted to be. I answered that I had two more years of middle school, four years of high school, four years of college and four more of graduate school before becoming a professor of physics.

XH: You liked to work on puzzles. How did you get into that? As an exercise for the brain?

SP: When we moved to the Bay area, the SF Chronicle had a Saturday puzzle page where I learned to do cryptic (British-style) crosswords. The page had other puzzles, including an algebra word problem. While growing up, my brother and I played lots of board games, including chess; but perhaps oddly enough I never took these too seriously and never became a real expert.

3. PROFESSIONAL CAREER

XH: Tell us what job placement was like back then. How did you decide to take a position at Harvard University?

SP: When I graduated from Stanford, statistics departments were growing all over, somewhat like the recent past. There were lots and lots of openings. So students, particularly from Stanford, had no trouble getting offers. I got a very early offer from Harvard. My interview was in December and the offer came in early January. It was likely I would get other offers, but Harvard wanted a reply from me before others could offer an interview.

XH: What was generally expected of a Ph.D. student? Did you have to have published papers?

SP: No. At that time, you were not expected to have published papers. There were very few postdocs, essentially none. Basically you should be finished with your thesis, but not always. Even if you were just finishing up, you could also get a decent position. That changed to some extent in the next few years. In mid 70s, there weren't as many openings. When I left Harvard, there were not as many really strong openings as I expected.

XH: So as long as you were almost done with your thesis. you were fine. It's mainly the letters from your professors that gave you jobs.

SP: Right. I am afraid of saying it this way, but it was the old boy network to some extent. I think statistics was always more open to women or minorities than most of the other areas. But at that point, most of the appointments would involve people on the committees knowing or hearing of your advisor. Your advisor's letter was very important, especially if he or she was very famous.

XH: So you didn't actually have to go to library to find the job ads.

SP: Not as much. No. In fact I only applied three places when I got my Ph.D. at Stanford.

XH: How did you know these three places that they had jobs?

SP: I think they were announced in the places like IMS Bulletin and AmStat News. But it's possible somebody at Stanford told me about one or more of them. So it was much more informal. It really was changing quickly at that time. By the time I left Harvard, I think it was already a requirement that you had to advertise an opening widely.

XH: What are the other places you applied to?

SP: I applied to Purdue and Berkeley. Berkeley was really looking for someone with a longer publishing record, so it was not likely. But when I told Purdue I had accepted Harvard (which was before they were making interview invitations), they seemed genuinely disappointed. I might add that Paul Holland (whom I had met while he was finishing at Stanford) alerted me to the fact that the 5-year appointment at Harvard had essentially no chance of becoming permanent. But statistics was a very fast growing field, and I expected that I could get good recommendations from very good people at Harvard.

XS: If I could ask, what's your starting salary at Harvard?

SP: I can tell you. It was $8000 a year. It was a time when the median salary in the country was five, maybe six thousand at most. I believe it wasn't that much different from now. About 1.5 times a median of over $60,000 will be over $90,000 salary today. Does that sound like roughly where many people get started?

But the advantage back then is that it did include full health coverage, and an additional 16% of salary went to your retirement. In fact, a couple of years after I started, they offered to pay some of the retirement benefit as salary. Also I should point out that the agreement with Harvard was to set up for a $1000 raise each year. So the starting salary was more like $10,000. This would clearly be similar to today's salaries adjusted for inflation in academic salaries.

XS: That's nice indeed.

SP: It was not a bad amount. One could live on that, sort of. In the Boston area, it was harder, but it could be done.

XS: Could you tell us a bit about your life at Harvard?

SP: At the time, Fred Mosteller, Bill Cochran and Art Dempster were the tenured faculty with a few junior Assistant Professors to provide enough course work for the degree programs. I had thought that Art Dempster might provide some interaction on decision theory, but he was much more interested in foundations and in applications. There really weren't any people with interests in more abstract mathematical statistics. Dick Dudley ran a seminar program in probability at MIT and people from several colleges in the area attended. I was the only one from Harvard. In 1972 (or so) Larry Brown had some NSF support to bring me to Cornell for a couple of weeks in the summer, and I had some great conversations with him and also got a chance to go mushroom hunting with Jack Kiefer and do some hiking with Peter Huber, who was also visiting Cornell. Thus, my continuing work in decision theory was mostly on my own with some crucial interaction via snail-mail (all there was then) with Larry Brown.

XS: What else influenced you at Harvard?

SP: Perhaps the most important contribution to my development was the arrival of Dave Hoaglin in my fourth year. He had been involved in the Princeton Robustness Study as a Ph.D. student, and gave a seminar course on Robust statistics. This really piqued my interest. I noticed that one standard sampling assumption that had not been subject to a robustness approach was the assumption of independent sampling. Actually, there was some work on testing problems by Herman Rubin and Ted Anderson (mainly in econometrics), but this topic was essentially untouched by the robustness field. Charles Stein had once suggested analyzing small perturbations to a parameter; and this led me to consider robustness to small amounts of auto-correlation spread throughout a sample. This led to my first nondecision theory paper, and gave me a great topic for a job-search talk a year later. In my opinion, this work still gives the best justification for using redescending score-functions in M-estimation.

The other great thing about Harvard was the quality of students. Harvard required that theses had a second reader, so I helped to direct thesis work of several great students. Of course, working with Persi Diaconis was great, but I also had great interactions with Sandy Zabell, and also with a mathematician, Gerald Edgar (who helped crucially with my results on admissibility of a one-dimensional coordinate of a multivariate location parameter). Nan Laird and Don Rubin were students at the time, and so given that there were only about a dozen Ph.D. students, the overall quality was remarkable. I also had my first Ph.D. student, Willis Davis, who was getting a Ph.D. while on leave from Lincoln Labs, and who returned to that position after getting the degree. Another nice feature of Harvard was the great mathematicians there. I established some contact and would often join them for sherry and cheese on Fridays. Andrew Gleason provided a crucial mathematical insight at one point, and I had some very interesting conversations with Sandy Zabell's advisor, Shlomo Sternberg. But not only were there no probabilists at Harvard, the conversations I had suggested that they knew much less about probability theory than I.

XS: When you moved to Illinois in 1974, was there already a research group in statistics at Illinois? Who were the key members at that time? Could you tell us a bit about your early days in Illinois?

SP: There was a research group. Bob Wijsman was clearly a leading statistician. And Jack Wolfowitz just arrived or was arriving. I don't remember whether he came the first year or the second year that I was there. But certainly with those two you have a strong research profile already. Bill Stout was moving into statistics already when I came. Bob Bohrer was doing a lot of applied stuff. The other person was Kumar Joagdev who was publishing papers very regularly on things like dependence and inequalities. Colin Blyth had been here, but left since I was filling his position. So there were about six people who were really called statisticians as supposed to, probabilists. But there were about six or so people in probability: Joe Doob and Don Burkholder were the best known, and I got to know Frank Knight, Walter Philipp, Catherine Dade, and later Ditlev Monrad well. I had extensive conversations with Walter Philipp when we had adjacent offices after we retired, and had some exchange of research ideas. There was a joint statistics and probability seminar. Within the math department, we had a sort of got-together as the probability-statistics group. So there was some organization to it. It wasn't quite a division. We had grants. At that time the grant situation was rather different. When I arrived, there were just two grants in the prob-stat group. As far as I know, everybody was on one or the other grant. That continued to be case until the early 80s. Then NSF began to split us off.

XS: How did you chair the statistics division in the math department for a couple of years before the Department of Statistics was established in 1985?

SP: After much reluctance from math, the department head was finally willing to let statistics go. Once that happened, the idea was to set up a division to start organizing things. The main thing I was doing then was trying to find a permanent head for a new department. There was sort of feeling that we need someone from outside to take over. Maybe that wasn't necessarily true. But it was one way to get an additional statistician in. We were given permission to look for someone to head the department. Basically, I chaired the division for one year, and then Jerry Sacks came to head the statistics department.

XS: What really led to the new department?

SP: I think the feeling was that math department did not really give us much support. For one thing, for example,

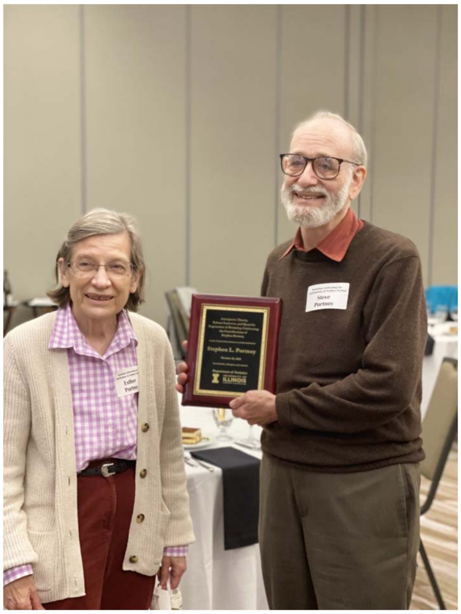

FIG. 4. *Steve Portnoy (with Esther) receiving the award plaque from the Statistics Department, University of Illinois at Urbana-Champaign.*

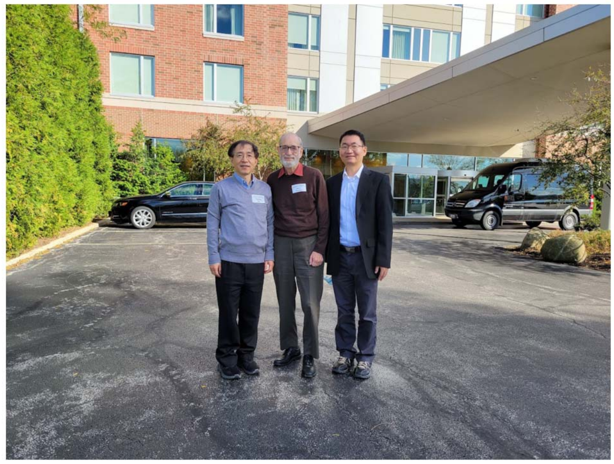

FIG. 5. *From left to right: Xuming, Steve and Xiaofeng in front of I-hotel after this interview.*

the overhead for our grants all went to the math department. They didn't give it back to us in any meaningful way. If we wanted our seminar speaker, we had to go to them to ask for support. Also it was hard to get Ph.D. students. At that time, there were a lot of statistics departments and if you were really interested in statistics, you did not apply to a math department. And the third thing is that the department didn't give any credit at all to applied work done with collaborators at other departments. The statisticians thought that collaborations were very important. Not just Bob Bohrer, but myself and others too. I think I was quite instrumental in trying set up the new department, and spent much time consulting with other people (especially outside the math department). As the first step, I organized an interdisciplinary committee in graduate school as we wanted to get support from other departments. This was a way to move forward and to point in the direction of a more general, more applied department. I mean previously the University of Illinois would have been thought of as highly theoretical with almost no applied statistics in a math department. That certainly changed, as was intended. We were hoping to get people who were more methodological and more applied. Over the years certainly, looking at the department today, that is certainly true. It's a very broad department now.

XS & XH: We are all appreciative of what you did (Figures 4 and 5).

XH: The field of statistics has always been evolving. What major changes did you see at UIUC statistics department prior to your retirement?

SP: I think that the changes in the department over time parallel developments in the field of statistics generally. One ongoing change is the (continuing) development of the field of statistics into a scientific discipline. At the risk of talking too much, let me elaborate: Before the middle of the last century, the ASA was primarily a professional organization, and (with the possible exception of Fisher) most statistical work had remarkably little serious mathematics. Slowly, academics began to take a greater role. The period following Neyman and Wald and the stars of the following decade (Stein, Kiefer, Robbins, Le Cam, Blackwell, Rao, etc) introduced a more solid scientific (and mathematical foundation). Most universities with statisticians developed separate departments. One clear indication of this development is that most papers in statistics now are jointly authored and have moderately large bibliographies (as is typical in most sciences).

When I started, almost all were singly authored, or if there was joint authorship, the paper was generally in a field of application with all other authors being nonstatisticians.

Also before the 1970s, it was rare for a senior statistician (of any kind) to have post docs and younger colleagues as part of a "lab" leading to publication of a moderately large number of papers developing a single main idea. I did not realize this when I first came to Illinois, but my own career certainly looked more "scientific" in time, with most of my subsequent papers jointly authored (with Roger Koenker, Jana Jureckova, Xuming He, and numerous applied collaborators); and for a while there was a "Quantile Lab" at Illinois—though it only consisted of 2–3 graduate students.

Another change is a greater acceptance of breadth within a department, especially as regards to applied work. Clearly, the UIUC department exhibits this increased breadth and orientation; and again my own career paralleled this: I always recognized the need for applied statistics to motivate theoretical research, but had only minimal contact with real data until after I came to

Illinois. My stint as director of the statistical consulting program introduced me to students and faculty in other fields. This led to papers in journals outside statistics: obvious ones like Theoretical Population Biology, but also Cell Differentiation among others, including work involving statistical analysis of the Hebrew Bible. Of course, my most important contact outside the department was Roger Koenker, who introduced me to regression quantiles.

A third change in our field is certainly the growth of the number of students. When I first came, I was the Undergrad Advisor to less than a dozen students, and the Master's degree was almost entirely a way-station on the way to the Ph.D.

Of course, there have been many area-specific trends within the field. Some, like MCMC and the AI approaches (data mining, big data) have become very hot topics, but I feel strongly that much more work, especially theoretical and mathematical, is needed before these approaches can be trusted. Specifically, it is critical to know exactly when and where each "black box" is likely to be effective; and I believe a lot more is needed along these lines, especially from the perspective of mathematical statistics.

XS: How did you start collaboration with Roger Koenker?

SP: He came to my office (around 1980 or so as I remember) and introduced me to regression quantiles. My robustness background helped me see the value, and I knew a little about linear programming (so I knew they were easily computable). The general value of a regression quantile analysis became quite clear as we talked, and our interactions grew substantially—eventually leading to a very long sequence of coffee meetings every morning (to which Xuming and others joined).

XH: Why did you take "early" retirement?

SP: Because of budget reductions, the University was encouraging early retirement for more highly paid senior faculty to free up funds for junior faculty. They offered a very nice 2-year raise, which significantly raised the pension. The pension also came with an automatic 3% raise each year (far more than we were generally getting). This provided much more time for research and graduate students, and made travel much easier.

XS: You become a co-editor of JASA (Theory & Methods) in 2005 after your retirement. Would you summarize your experience as editor? What did you try to achieve at JASA?

SP: I believe Walt Piegorsch and I were the first joint editors of JASA-T&M. I had been an AE for the T&M section for a rather long time, and I think I had convinced several of the editors that I could provide excellent AE reports that matched what they wanted: even T&M papers should show a clear applied motivation. So, despite my reputation as a theorist, the ASA Board was willing to appoint me as a co-editor. In fact, Walt and I simply divided the papers by each taking every other submission. One problem with this is that since each co-editor was seeing only half the papers, we didn't notice that we were rejecting a larger than usual fraction of the submissions for some time. A positive feature of this is that we eliminated a long backlog, but it wasn't easy to increase the acceptance rate back to a bit over 25%. One of our main objectives was to expedite the reviews. One major problem was that (unlike most reviews in science journals) stat reviews tended to nitpick details and ask for very extensive revisions. I think this is a consequence of academic statistics coming from math, but it was very hard to overcome the inertia (especially since almost all reviewers had their own submissions nitpicked!). We tried to keep close tabs on the progress of all AEs and encourage them to bug referees to get a quick report when the review process was dragging on. But the most effective idea (Walt's) was to get AE volunteer to write a referee report on a paper that had been in the pipeline for 4 months or more. Quite a few AEs were willing to do this; and the fact that this could be done often pried loose a report from the original referee. We did manage to reduce review time considerably (in terms of the mean and also by a censored quantile analysis).

Otherwise the main objective was to maintain quality; which I believe we did. One highlight was that my tenure included the 100th anniversary of Student's paper. I got Sandy Zabell to write a very nice historical analysis of the paper [5], and got several other well-known statisticians to discuss it. In general however, both Walt and I tried to keep discussion papers at a minimum.

I really enjoyed the editorship, and, of course, it was a great honor. But it was a huge amount of work (even as a co-editor), and I can't see how anyone can do this without some slowing of research activity. Fortunately, I had taken early retirement by the time, so had no teaching duties.

XH: Tell us who have been most influential in your research career in statistics. What advice do you have to younger statisticians today?

SP: Certainly Charles Stein was very influential in many many ways. The advice is: don't be afraid of asking someone who is very famous, or who might be very mathematical, or very applied if that's your interest. I think it's always good advice to think hard and look for who can really help you.

After that, the next person who was very influential was Roger Koenker. As I noted earlier, Dave Hoaglin introduced me to robust statistics. Lots of people have contributed to the ideas that led to the research that I did. Certainly advice from others and collaborations are absolutely essential. So be willing to talk to others, be willing to ask questions and be willing to offer help when needed.

XH: That's very valuable advice, Steve. Even more valuable today. If you just work on your own all the time, it's not going to work.

SP: I think this is one of the things in statistics that is pretty indicative we have moved from a pure mathematical area where people did almost all their work on their own, to a more scientific type of environment where there was a lot of collaboration, a lot of joint authorship. There are postdocs. There is a big change in the way that statistics operates now on the academic level, compared to when I started. Academic stat looked much more like math when I started. Papers had very short bibliographies. (all laugh). I don't know if mathematics is still that way as much. But I think it is still more that way than statistics.

XH: You were awarded Francqui Professorship in Belgium. Few statisticians have received this honor. Can you tell us how it was like?

SP: Marc Hallin is a very well-respected researcher in Belgium, and he is very good at submitting successful proposals to the Francqui board. As I remember, Francqui was an engineer who had worked with Herbert Hoover in the early 20th century. He provided funding to establish a national program to bring eminent scientists to Belgium for a semester or two, and so the vast majority of Francqui professorships were not in stat or prob. Of course, it was a great honor. It was a great deal of fun. I had only one course to teach: a once a week seminar course on regression quantiles, and I was required to arrange a 1-day workshop, which turned out to be quite successful. Being in Europe, it was relatively easy to visit several colleagues there. However, I was JASA editor at the time (thanks to being able to do the job entirely online), and I had hoped to have more time to engage in joint research with Marc and other statisticians there, especially Mia Hubert whose student was finishing her thesis following some joint work we did. Esther was also retired at the time, and her job was to try all the chocolate shops in Brussels (of course, an impossible task). We did get to make a number of friends in the 6 months, and feel we got an opportunity to see how continental societies worked, at least from the point of view of an academic.

XS: You have been a mentor for many Ph.D. students. Can you share some of your fond memories of the students?

SP: Students keep you young. They are a great source of new ideas and keep you on your toes. I certainly remember right from the first student I've had at Harvard to the last one I just had a few years before leaving Illinois. Some of them were really quite strong mathematically and could do almost everything on their own. It was more like working with colleagues. It was certainly true with Xuming, you might remember. We talked about a lot of things without my saying "Do this, Do that," or my having to offer very much in the way of nudges. Some of the students required a great deal of more in the way laying out how to go about the next step, what to read, things of that sort. But usually having them once a week to talk about what they are doing was very important. And I think I got to know them as friends as well as colleagues. Actually, I directed only about a dozen Ph.D. theses, and I remember all of the students fondly.

And let me add this point, I was sort of a second reader on quite a few Ph.D. committees in the biology department here. And that was a lot of fun. There, they didn't know as much statistics. So I had to give a lot more detailed ideas on how to go. But you know it's a lot of fun to see how the statistical ideas get applied and how they work out. I felt I was doing a great deal of good by being on these committees and working with these students. And again it was a very enjoyable experience.

4. KEY RESEARCH AREAS

XH: You are among the earliest to work on statistical problems with the parameter dimension increasing with the sample size. What motivated your work there? What do you think about the more recent boom on high dimensional data analysis?

SP: The modern question is really rather different from what motivated me. I think some comments in Huber's book may have been the initial catalyst, but I also recognized the increasing use of normal approximations in moderately large dimensional problems—and it was obvious that if the dimension, p, wasn't small compared to the sample size, n, then the multivariate normal approximation couldn't really be good. So my goal was to try to quantify this in some specific areas (e.g., M-estimation). Basically, to get uniform normal approximations in p dimensional parameter spaces, you may get by with $p^{3/2}/n \to 0$ (if there is sufficient symmetry), but generally you will need at least $p^2/n \to 0$ (and there are examples showing that you really can't let p grow more quickly in general). Thus, the modern paradigm with p growing often exponentially faster than n is clearly a very different situation.

Generally, what saves the day now is the assumption that there is really a fixed dimensional subspace and one seeks some sort of oracle theorem. Also, there has been some work looking at cases where p/n is roughly constant, eg, the eigenvalues of a Wishart matrix, and there was some quite intriguing work by El Karoui, Peter Bickel, Bin Yu (and others) several years ago. But clearly letting p grow more quickly than n is going to require very stringent conditions, and really must be considering something other than uniform normal approximation in p dimensions. The kinds of conditions are going to depend strongly on the specific application and question; and I doubt that there is a single approach to large p problems.

XS: You and Roger Koenker have been collaborating for many years on topics in robust statistics and quantile regression. Could you tell us how the collaboration worked?

FIG. 6. *From left to right*: *Xuming He*, *Roger Koenker and Steve Portnoy*; *fun time in Urbana*.

SP: It got started when he came to my office as I mentioned. But it continued quite a bit. There was a long period of time when we met daily for coffee. You remember those, Xuming?

XH: Yeah. The memorable coffee hours at Espresso Royale, and more (Figure 6).

SP: I am not sure when the coffee meetings began, but I believe they started with just Roger and me in the 1980s. Xuming joined when he returned to Illinois as faculty in the 1990s, and they became much more regular. I had a series of good Ph.D. students from the 1980s–1990s: Lin-An Chan, Xuming, Liji Shen, Kenneth Zhou, Quanshui Zhao; but I don't remember any but Xuming coming. Later (after 2000) students of both Roger and me came more often (e.g., my student Teresa Neocleous) and some junior faculty (I remember Naveen Narisetty coming regularly after he arrived).

As we got a little bit more famous, people would be coming to us from all over, from Europe, from other places. Very often they would join us for coffee, and be working with us. In fact, some were people who were getting their thesis from somewhere else and came to us for 6 months or a year. That was a lot of fun. They came mostly from Europe. Roger mentored several such students, but two who I mentored were Arne Bathke (a student of Manfred Denker in Goettigen) and Karlien Vanden Branden (Mia Huber's student in Leuven).

XS: My collaborator, Stanislav Volgushev, was one of the visitors you brought from Germany, right?

SP: Yes, he came to work with both Roger and me.

XS: Thank you for doing that. We had already written five papers together. (all laugh)

By the way, do you want to say a little bit more how this morning coffee meeting worked? Did you have a very structured way of what's going to happen tomorrow morning?

SP: Usually not. We often met at 8:30, relatively early in the morning. We would meet, and there used to be some informal discussion. The weather was always there. But the things we were working on would come up. Very often we went on for several weeks on a single topic. Other times it would be changing, especially when someone was visiting. They would often join us for coffee and so the discussion would shift to something they were interested in. It was very informal, but there was a lot of discussion in statistics and mathematics. Sometimes we covered napkins with all kinds of equations.

XS: Proving a theorem on napkins? (all laugh)

SP: At least it was the basic idea.

XH: What do you think is the most important contribution you have made in this area of quantile regression?

SP: Like most scientists, most of my work has provided incremental improvements. That is, my colleagues and I have developed new methods and approaches that can be expected to work in cases where the data are more complex than had previously been assumed. I have contributed a few theoretical results that I feel are quite important and fundamental. Two examples in quantile regression: (1) the number of regression quantile breakpoints is roughly of order n, the sample size [1], and (2) a Kaplan–Meier type result replacing the $n^{-1/4}$ normal approximation error from an almost sure Bahadur representation by a $\log(n)^a n^{-1/2}$ error rate; see [3]. There are two ideas I contributed to quantile regression that I feel are really revolutionary: (1) the fact that regression quantile computation can be expected to be faster than the least squares in typical regression problems with large n [4], and (2) censored regression quantile analysis [2].

XH: Robust statistics was an active area of research 40 years ago. Where do you think robustness needs to play a bigger role in data science today?

SP: I would say lots of places. One would be big data and AI areas. Much work here is based on certain black boxes and complicated methods which have not been carefully analyzed as mathematical statisticians would. It is often not clear when you first start what the model really is, or where it comes from. But now I think paying attention to where it comes from and thinking about what happens if there's a perturbation of the basic assumptions that you make is critical. I think the idea of doing that by looking at perturbation theory, which is what robustness is about, is still there and simply hasn't been applied in these areas very much. I think it could also be applied very usefully in Bayesian analysis. It should be possible to look at what happens when you change the prior just a small amount and what happens when you changed the model just a small amount and try to look for minimax type solutions. As far as I know, I just don't see the Huber (minimax) approach being tried, and I'm not sure why.

XH: Maybe it is just very difficult, you know, difficult to formally quantify effects and prove specific results about robustness. In Bayesian statistics, at least, there have been efforts to talk about robustness of the prior.

SP: Yes, there has been some work, though I have seen very little aimed at finding simple ways to adjust for very general departures from assumptions. Most seem to be aimed at "continuity" arguments showing that if the departures are sufficiently small, you don't have to worry very much; or they just suggest enlarging the models.

XH: Speaking of robustness to assumptions, I want to ask a related question. You talked about the machine learning, and black box type of things, but one might argue that some machine learning methods really are not making any assumptions on the model, so why do you still think that robustness is still an issue?

SP: If you give me a black box, I would claim there's a model that underlies it and all you've done is cover it up. If black box works, it's going to work for only certain times, or certain kinds of problems. Now if you don't want to write it down that's fine. But in my opinion, it's just impossible to think about fitting a structure to data if the data set is very large. Mainly, there are far too many structures, far more than the number of atoms in the universe. I am actually sure that anyone of us could, given a large data set, find dozens of models just as good as anything given by a black box. Basically, just partition the data and make incremental improvements in each subset.

XH: So you are thinking about and interpreting models very broadly. Not the kind of probability models or parametric models.

SP: They didn't have to be, but I think that some mathematical form would be necessary. And probably if you're thinking about the things that are random, in some sense that makes sense to consider models which have a probabilistic component. But I would think the engineer would be very conversant with differential equations and shouldn't have any trouble figuring out models underlying most of these computational methods. They might be rather complicated but one should be doing it because that's how you can see what is really going on. Mathematical statistics could really make some progress on what is really being assumed and when it's likely to work and when it's not likely to work.

XS: I noted that you have also done interdisciplinary research with some biologists at UIUC. How did that happen? Can you tell us your favorite piece of work outside theoretical statistics?

SP: It happened in the first few years I came here. Bob Bohrer went on sabbatical, and I was asked to take over this consulting program. It was really a one man job at the time, and one of the people who came in was a student (Steve Ferris) who had this very interesting data set on catfish. He had thirty species of catfish. The catfish family had originated about 55 million years ago by duplicating its genome. So the progenitor had twice as many genes. Instead of having just two copies, each of them had four. So these genes could evolve rather quickly because one didn't lose the original genes when the new ones changed. So what would happen over time is that, some of the duplicate genes would become dysfunctional (silenced), but many of them would now exist in two forms (i.e., 4 copies). He and his advisor Greg Whitt noticed that the more advanced the catfish was (i.e., the more speciation events in its past), the more genes became silent. So he did a regression of the number of silenced genes versus the number of speciation events, and got a very nice straight line and did a black box : very significant. They submitted a paper, but the referee said: but you've got an evolutionary tree here which means closely related species are highly dependent. And you haven't included dependence in your regression. The student (and his advisor) needed my help.

This is a long answer, I know, but basically I said, well, there are some techniques assuming dependence, but you have a lot of information here. You have this evolutionary tree which you can actually write down, and what I did was working with them to develop a birth-death process along the evolutionary tree and solved it as a statistical problem in applied stochastic processes. It took a lot of work. The likelihood that came out at the end was a product of something like 26 polynomials of up to degree six. So it was very complicated to solve those days. But it was a lot of fun. That was certainly one of the more interesting ones. After that I got to know some of the people of the biology and as students came along who either had a statistical problem on their thesis, or the thesis was a little bit more mathematical, they would ask me to be on the thesis committee. So I was on the thesis committee for about eight students. Most interactions led to very nice papers in biological journals, and I would often get my name on a paper this way. Because it appeared in the biological journals where bibliographies tend to be much longer, they would get lots of citations and so some of my papers that are most cited are not in statistics.

XS: Yeah I think that's really true.

SP: I think anyone who does applied work knows that. It is something that is changing. Our bibliographies are getting longer, and I think statisticians are getting more citations now, than they did when I first graduated. The field is becoming a little bit more scientific.

5. REFLECTION

XS: In your opinion, what have been the major breakthroughs in the field of statistics in the last 50 years? Where do you think the field of statistics going?

SP: The basis of a breakthrough is of course a highly innovative (revolutionary) approach to a problem. But I don't consider any new idea as a breakthrough until it has a thorough theoretical underpinning AND a very large amount of successful empirical experience. Obviously, I would point to Stein's rule as a breakthrough. But note

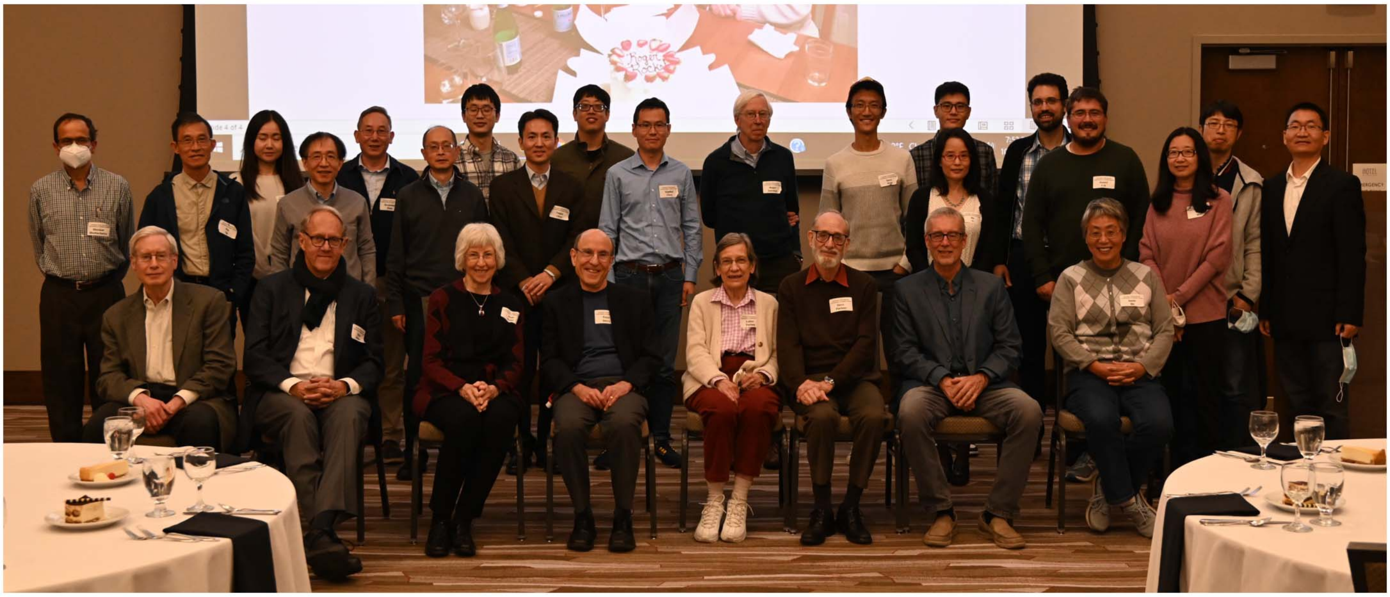

FIG. 7. *Conference Dinner at the University of Illinois Workshop in honor of Steve Portnoy on October* 16, 2021.

that the original result was in a 1955 Berkeley Symposium paper showing that a very small perturbation of the best invariant estimate of a multivariate location parameter would give a small but uniform improvement in mean squared error (even without normality). The James–Stein paper five years later finally provided evidence that the improvement could be useful, and the subsequent extensive development (and later connection to regularization) is what made this a breakthrough. The list of contributors to this is far too long to give in detail, but certainly included people like Larry Brown, Bill Strawderman, and Mary Ellen Bock, who did a thesis here mostly directed by the econometrician George Judge.

Another breakthrough is certainly the bootstrap and related resampling methods. I would add one more: the attention to what happens when the classical finite parameter assumptions fail. We now have relatively well-developed methods and theory on robustness and on non- and semiparametric models. Of course there are lots of more specific breakthroughs in specific areas. Quantile analysis in regression problems is certainly one. I do think MCMC is a breakthrough in certain specific problems, and things like the analysis of false discovery rate have useful applications; but I have a feeling that many of the extravagant claims for some of these "breakthroughs" are decidedly premature. Fortunately, the future of statistics is in the hands of the next generations of statisticians, and the ratio of highly innovative work to incremental improvements seems to be at least as large as it has ever been.

XS: Are there any experiences from your career that you'd like to share with today's students interested in academia as well as our junior faculty?

SP: I did mention a couple of things, like not being afraid to pick a really good thesis advisor. Collaboration is important, and if you are shy and a little bit afraid, get some close friends to help you talk to other people if you need help. Don't be afraid to ask very famous people. They are usually more than happy to talk to you. Certainly I benefited greatly from being willing to talk to whoever I could.

I think that statistical science will continue to be most successful when it focuses on drawing inferences from a sample to a population, Specifically, it will always be critical to delineate in what situations specific methods can be expected to be useful. That is, in terms of modern big data analysis, it is critical to know when a specific black box is likely not to be stupid. Of course, statisticians (and many others) will continue to analyze data, sometimes usefully and sometimes not; and distinguishing this will require continued statistical science.

I expect that statistics departments will continue to be attractive and well regarded for some time, and will continue to provide mathematical study on the analysis of statistical procedures, continued development of methodology, and ample opportunities for collaborative data analysis. I wish my younger colleagues a great deal of joy in pursuing these goals

XS and XH: Thank you, Steve, for sharing your wisdom and enthusiasm with us today (Figure 7).

REFERENCES

[1] PORTNOY, S. (1991). Asymptotic behavior of the number of regression quantile breakpoints. *SIAM J. Sci. Statist. Comput.* **12** 867–883. MR1102413 https://doi.org/10.1137/0912047

[2] PORTNOY, S. (2003). Censored regression quantiles. *J. Amer. Statist. Assoc.* **98** 1001–1012. MR2041488 https://doi.org/10.1198/016214503000000954

[3] PORTNOY, S. (2012). Nearly root-n approximation for regression quantile processes. *Ann. Statist.* **40** 1714–1736. MR3015041 https://doi.org/10.1214/12-AOS1021

[4] PORTNOY, S. and KOENKER, R. (1997). The Gaussian hare and the Laplacian tortoise: Computability of squared-error versus absolute-error estimators. *Statist. Sci.* **12** 279–300. With comments by Ronald A. Thisted and M. R. Osborne and a rejoinder by the authors. MR1619189 https://doi.org/10.1214/ss/1030037960

[5] ZABELL, S. L. (2008). On Student's 1908 article "The probable error of a mean." *J. Amer. Statist. Assoc.* **103** 1–20. With comments and a rejoinder by the author. MR2394634 https://doi.org/10.1198/016214508000000030