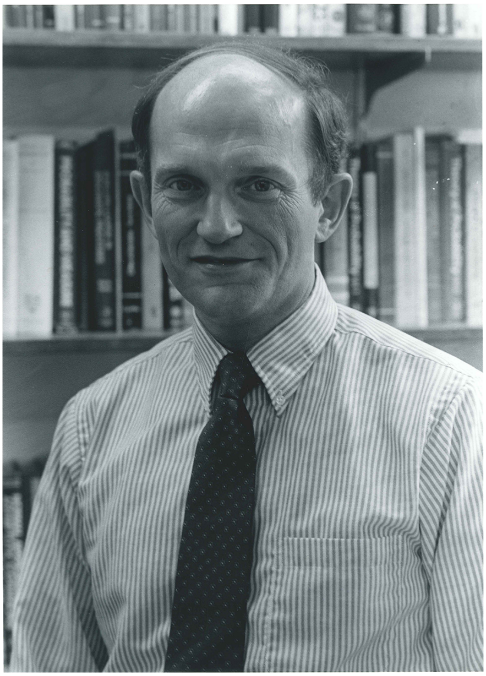
Econometric Theory, **39**, 2023, 1–26.
doi:10.1017/S0266466621000372

THE ET INTERVIEW: PROFESSOR GARY CHAMBERLAIN

Interviewed by Bryan Graham, Keisuke Hirano, and Guido Imbens



Gary Chamberlain was one of the most important thinkers in modern econometrics. He wrote deeply influential papers on panel data analysis, latent variable and qualitative response models, semiparametric estimation, quantile regression, asset pricing, and decision-theoretic methods in econometrics.

Equally important to his written output was his impact on his students and colleagues. He was profoundly generous with his time and his insights, and his influence can be discerned in numerous papers over the past few decades of research.

Gary Chamberlain was born on April 23, 1948 in Boston, Massachusetts. He attended high school at Boston Latin School, and started college at Harvard in 1966. Upon graduation in 1970, he entered the PhD program in economics at

Bryan Graham: Department of Economics, University of California, Berkeley, California, USA
Keisuke Hirano: Department of Economics, Pennsylvania State University, Pennsylvania, USA
Gudio Imbens: Graduate School of Business, Stanford University, California, USA; email: imbens@stanford.edu

Harvard and finished his dissertation under Zvi Griliches in 1975. His first job was as an Assistant Professor in Economics at Harvard University. In 1979, he went to the University of Wisconsin-Madison. In 1987, he returned to the Harvard Economics department where he remained for the rest of his career. Gary made seminal contributions to many areas in econometrics, including the econometric analysis of panel data, decision-theoretic approaches to econometrics, Bayesian econometrics, and semiparametric models. He advised many students and was an influential role model and mentor for many junior econometricians. He was a fellow of the Econometric Society, the American Academy of Arts and Sciences, and the National Academy of Sciences, and was elected as a Distinguished Fellow of the American Economic Association. Gary served as Co-editor of Econometrica from 1983 to 1986.

On January 3, 2016, Gary Chamberlain sat down with Bryan Graham, Keisuke Hirano, and Guido Imbens at Bryan Graham's home in Berkeley, California for an extended conversation about his life and his career as an econometrician. The day before, Gary had been honored as a Distinguished Fellow of the American Economic Association at the ASSA Meetings in San Francisco, along with Ted Bergstrom, Tom Rothenberg, and Hal Varian. The questions were asked by Guido Imbens (GI), Bryan Graham (BG), and Kei Hirano (KH).

Gary Chamberlain passed away unexpectedly in February 2020. It was a huge loss to the international community of scholars in econometrics, but Gary's influence will undoubtedly continue to be felt through his research and his many students, colleagues, and friends who have been inspired by his example. In his honor, the Gary Chamberlain online seminar was started in the spring of 2020.

GI: We thought it might be good to do a little bit chronologically first. Would you like to say anything about high school? Where did you go to high school?

GC: Boston Latin. Starting more at the beginning, my parents grew up on farms in New Brunswick, Canada. They knew each other, but I think they actually started dating in Boston, where I was born. As far as I know, neither had finished high school, which I think has pluses and minuses, but on the whole, it has been a plus for me. I never felt like I had to surpass my dad, whereas I think Rachel and I have given our kids a certain burden because we are academics. We lived in downtown Boston, very close to Boston Latin. So that was very good when I got admitted. I started there in seventh grade; for the first six grades, I remember going to Mission Hill School, which was in a dangerous area. At Boston Latin, much of what I remember is rote memorization. It was a very competitive place; I got into that. Before that, I actually enjoyed learning on my own; I was interested in science and that was fun. But at least initially at Boston Latin, it was competitive—just get good grades. It was not a whole lot of love of learning. I got into that and did okay but for some reason, in sophomore year in high school, so 10th grade, I do not quite remember why but I just sort of stopped feeling like I wanted to do that. I did some rowing; there was a crew team. I had a really big slump in my grades, which was good, because then I picked up in junior year and got back into things and got the most improved award at graduation which had some substantial money attached to it, which was appreciated. But in terms of actually enjoying learning in high school…

I have a very clear memory that between junior and senior years, I went to Harvard summer school, which had a program for high school kids. There was one other kid from Boston Latin, Howie Possack. We were there together and we were pretty good friends then, and it was a calculus class taught by Clifford Earle, a real mathematician from Cornell and it just made a huge impact. I just loved it. I thought, this is magic. I went back to Boston Latin for senior year and there was an AP Calculus course taught by Mr. Branckert, I think. He relied on notes that apparently were very, very old, and it was a bit of a tradition for people to try to steal the notes. There would be periods where he would just be at sea. It was good that I had the summer experience, because I knew that calculus was good stuff, but his course was not inspiring.

I guess the other nice thing about the Harvard summer school is that we got to use the rowing shells. You could start with a broad comp, narrow comp,[1] and eventually get into a shell, and that was really a revelation because the seat moved. The crew at Boston Latin would compete against Boston English, Boston Tech, Boston Trade, but those boats had fixed seats, they were sort of large dinghies really, and you would get splinters in your ass from them. Whereas those moveable seats, those were great.

For college, I kind of thought I would go to West Point: it was free. And I got into West Point, but then I think partly the experience at Harvard summer school and the people there that suggested I might consider other things. I did apply to Harvard and MIT and got into those, and decided to go to Harvard.

Okay, back to Boston Latin: Chuck Manski is also a Boston Latin graduate. I do not remember a lot of interaction [with him] during that period. It was really at Wisconsin where that occurred. Although I think Possack would be a bit of a connection, because I think he was quite friendly with Chuck and I got quite friendly with him [Howie Possack].

GI: You guys were in the same year?

GC: Yeah, we were in the same year. But not much interaction really before Madison. I remember on the job market being interviewed by Carnegie Mellon and I think Chuck was in the room. I think he was at Carnegie at that time. But no, it was really at Wisconsin where I got to know Chuck.

GI: So, when you went to Harvard, you were an Economics major. How did that go? (Figure 1)

GC: I was. Again because of this calculus course, I was thinking, well maybe I would do math, and I did a certain amount of reading on my own the summer before Harvard. If you indicate an interest in math, they have got a kind of placement mechanism. And I somehow ended up in the office of Andrew Gleason, a distinguished mathematician. And I was telling him how over the summer I had

[1] "Broad comp" and "narrow comp" refer to two types of boat used at Harvard's Weld Boathouse, where "comp" is short for "compromise" between a wherry and a single.

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

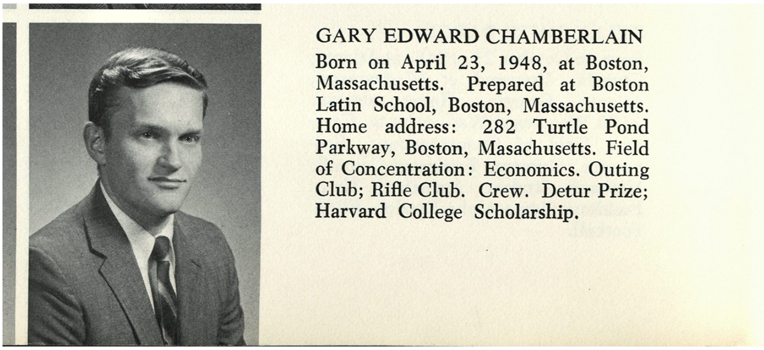


FIGURE 1. Harvard College Yearbook, circa 1970.

been studying "Lo Bes Goo" integration. He was very interested initially until he realized "oh you mean Lebesgue?" He was really more interested [in questions like] do you like to do mathematical puzzles? Basically, the answer to many of his questions was no, so I was really being tracked out of the pure math line. I took a fair amount of math and enjoyed it. I guess in terms of thinking about influential courses at Harvard, Bradley Efron was teaching upper-level stats, an undergraduate course. I am not sure if DeGroot was the book, but that was the level, and I still remember the enthusiasm of his lectures. It really had a big impact. Very different but also a big impact was Arthur Dempster. His lecturing style was very different, and I suspect was not the sort of style that would generate good ratings. But it just made a huge impact. That was a graduate course based on his book *Elements of Continuous Multivariate Analysis*.[2] And yeah, there was a period when that was a big part of my life, trying to figure out what was going on in this book. He was somewhat helpful with his class lectures and he was great to sit down and talk to. But it had a big impact because I would look at other multivariate analysis books and there would be all these formulas of bizarre densities, formula after formula. I think Arthur's book maybe has no density formulas. He was just able to do everything by setting things up in a coordinate-free way, in an abstract vector space and he could get all the results just from geometric reasoning. Huge impact, huge impression.

Other things as an undergrad were very influential: Marty Feldstein hired me as a research assistant, probably by sophomore year. He would pretty much just suggest things to work on. Okay, he suggested, there was this paper by Balestra and Nerlove on dynamic panel data,[3] and I think he was interested in actually using it for something, so he wanted me to program stuff and think about it. That had

[2]Dempster, A.P. (1969) *Elements of Continuous Multivariate Analysis*. Addison-Wesley.

[3]Likely Balestra, P. & Nerlove, M. (1966) Pooling cross-section and time series data in the estimation of dynamic models: The demand for natural gas. *Econometrica* 34, 585–612.

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

some impact. But it was really more just a sense that being an academic might be something I would actually like, and I think he was the person that I got the most sense of what it was like to be a professor.

GI: What was it like, on a daily basis, to be a research assistant in those days? You were doing some programming as well as some reading?

GC: We had a terminal attached. Otto Eckstein was running DRI[4] at the time, and there was a computer hookup via that to a terminal. So yeah, a good bit of it was computational work.

GI: Where was the computer, the terminal, actually located? Was there a sort of office where there were some research assistants?

GC: Marty had kind of a suite so there was an office right next to his office where there would be two or three of us often using it. I would use it at night a lot. But a lot of work also was just across the street, and at this point, I guess we were in 1737 Cambridge Street. Across the street there was an IBM mainframe, and that was a punch card operation. Very relevant, even relevant for the paper I am presenting on Monday,[5] Marty had me working on mean-squared error improvements. So, Dudley Wallace would be probably one of the bigger influences that economists were aware of, the work that he had done. Marty's setup was very straightforward. Normal linear regression with two variables, you care about β. X_2 is there to avoid bias, but maybe for mean-squared error you could do better by not doing the least squares fit on X_1 and X_2. And yeah, from Dudley Wallace and various things statistics people were aware of, Marty wanted to see what the mean-squared error of various things was. Like drop X_2 based on a t-test. Use X_1 and X_2, but take a weighted average of the coefficients based on a t-squared over one plus t-squared, things like that. And I would program simulations to calculate mean-squared error curves. That was published in *Econometrica*.[6] Not sure, but it may be the most straight econometric piece that Marty has published. Shrinkage stuff is still of interest.

GI: It also seems reasonably close to some of the things Ed Leamer was doing later? Was there a connection? Leamer was not there yet, right?

GC: No, this was sophomore or junior year. Leamer was there when I was a graduate student. But Chris Sims was there, and I do recall Chris saying, well you should think about a Bayesian interpretation of this; what sort of priors would result in this or that. Particularly the notion of using both variables but with appropriate shrinkage and a Bayesian interpretation of that. I recall Chris talking about that and pushing things a bit in that direction.

Chris Sims was my undergraduate thesis advisor. I think the thesis was good to have done; I do not think there was anything particularly memorable there. But

[4]Data Resources, Incorporated (DRI) was founded by Otto Eckstein and Donald B. Marron. See Arenson, K.W. (1984) Otto Eckstein, Educator Who Led in Economic Forecasting. *The New York Times*.

[5]Chamberlain, G. (2016) Fixed effects, invariance, and spatial variation in intergenerational mobility. *American Economic Review: Papers & Proceedings* 106, 400–404.

[6]Feldstein, M. (1973) Multicollinearity and the mean square error of alternative estimators. *Econometrica* 41, 337–346.

talking to Chris was like talking to Marty: there was a sense that there might be a life here that would appeal to me. And Chris was very open-minded in terms of if you had some interest in kind of a math question, he would say, well go ahead, this is a good time to do that. I do recall in particular… well there are probably two things. On the times series part Rozanov's time series spectral analysis book,[7] it was really more on the probability part not so much on inference, but yeah that is a book I tried to get on top of, with encouragement from Chris. Another book was [*An Introduction to*] *Harmonic Analysis*, by Katznelson,[8] that one, I remember, that is something I have come back to later. But that was sort of in the background for quite a while; at the time I thought yeah, I would like to know more about this.

I guess, at the end of Harvard, the summer after graduation, Marc Nerlove was running a workshop. It was Marty who told me about that. It was supposed to be for graduate students, but Marty said, what are you doing for the summer? This would be my senior year; it would have been some sort of RA work. The summer after freshman year I actually spent changing light bulbs in the Harvard Med School labs, from which you learn that there are risks in experimental work. You are above those [animal] cages and you have got to watch out because they can reach up. But anyway, I did not have really specific plans. And he said there is this program that Marc Nerlove is running; it sounds like it might fit you, and he contacted Marc Nerlove and I got admitted and got invited. I was the only nongraduate student, everybody else was working on their theses. Mark was then at the University of Chicago, and he was running this workshop jointly with G.S. Maddala and David Grether. They would have eminent economists or econometricians come through and give talks. And what were we doing? Well, people would talk about their thesis work. I guess I talked a little bit about my senior thesis and G.S. Maddala told me that it is good you did this, but you should forget about it. And a lot of interaction with G.S. Maddala, Nerlove, and Grether. So that was interesting. Zvi Griliches passed through and gave a talk. The setup would usually be that you would all go to dinner with the visiting famous person. Okay, I am remembering a bit more now. Marc had to give one of these, it was something like the Fisher-Schulz lecture—it was probably the Fisher-Schulz lecture.[9] Because I think he became the president of the [Econometric] Society later, and he would have been working on a talk for that. He was basically spending the summer working up one of these lectures like Fischer-Schulz, and very interested in economic dynamics and implications for econometrics and dynamic programming. So, he kind of arranged to have people come in that might be relevant, so that was part of what led to this. That is when I met Zvi, who was there for at least a day.

[7] Authors' note: we believe Gary is referring to Rozanov, Y.A. (1967) *Stationary Random Processes*. Holden-Day.

[8] Katznelson, Y. (1968) *An Introduction to Harmonic Analysis*. Wiley.

[9] Authors' note: Nerlove gave the Econometric Society's Fisher-Schultz Lecture in 1970, which coincides with the year Gary completed his undergraduate studies. A revised version of Nerlove's lecture appeared as Nerlove, M. (1972) Lags in economic behavior. *Econometrica* 40, 221–251. Based on the acknowledgments in the lecture, the summer workshop Gary is recalling was the Mathematical Social Science Board Workshop on Lags in Economic Behavior held at the University of Chicago in 1970.

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

GI: So Zvi was not at Harvard yet?

GC: No, I think he was on his way. I think when I showed up in my first year of graduate school, it was maybe when he showed up—pretty close. So, I guess now we are at the point where I am a graduate student.

GI: How did you decide to go to Harvard? Was that sort of the obvious thing to do?

GC: I had conversations with people. In retrospect I guess when I have these conversations with a student [now], I tend to push them a bit away [push them to go elsewhere]. I would say I was kind of encouraged to stay. MIT, I guess, was the other option.

GI: One more question about the program Marc Nerlove ran. Were any other graduate students there?

GC: Peter Schmidt was there. And some Australians, but in particular, Grant Hillier. Which is interesting because, much later, I worked on stuff that at some point I realized was quite related. The kind of canonical forms I would set up for IV, Grant had done very nice work [on that] before.

GI: The first year in graduate school, who was teaching the econometrics sequence? Henk Houthakker must have been teaching part of it?

GC: No, I do not think Henk taught the econometrics sequence. He would go around saying that econometrics is a young man's game. Who was teaching? I did not take many courses. In my first year, I was a teaching fellow for Zvi and possibly also for Dale, so I have a total blank on what the first year was.

GI: Zvi had been there longer already?

GC: Little unclear, we should check. But Zvi and Dale arrived very close, maybe simultaneously, and Ken Arrow I think likewise.[10] There was a big young influx right around the same time. I do remember Ken Arrow teaching the first-year theory course. I took that and that was quite influential.

GI: Was there a sense that the department was really rejuvenating at that time?

GC: Yes. I was an undergraduate becoming a first-year graduate, so I was not plugged in. But yes, I think there was a sense that they needed to do something. And the arrival of Zvi, Dale, and Ken Arrow was something.

GI: And there was a sense among the graduate students as well? Obviously, it was much less coordination, nowadays, I think the students would all know this is happening.

GC: Yes, it was so different then. My decisions of where to go to graduate school, did I know the major players were here and here? No. That is just so different now. And having a recruiting day today…

GI: Who did you end up talking to the most as a graduate student?

GC: Ed Leamer showed up quite soon, maybe my first year or second year, and that became a big deal.[11] We would talk a lot; two [co-authored] papers came out

[10] Authors' note: Zvi Griliches moved from Chicago to Harvard in 1969; Dale Jorgenson moved from the University of California-Berkeley that same year; Ken Arrow had arrived the year before, from Stanford.

[11] Authors' note: Edward Leamer arrived in the summer of 1970, from Wayne State University.

of that.[12] His book was in progress and it is got to be up there with [my] the major formative influences. It just seemed to make a lot of sense.[13] Chris Sims certainly was very open to Bayesian thinking. Marty Feldstein I guess likewise, apparently a large influence on him was Terence Gorman, who I guess very much presented things in a Bayesian way. The discussion with Ed in terms of how to think about stuff was very influential. But the other big thing would be Zvi Griliches; he was the chair of my thesis committee. The committee was Zvi, Marty, and Ed Leamer. So Zvi was very interested in some panel data issues that I guess were somewhat cutting edge at that time. He was very interested in work Art Goldberger was doing, so [he] plugged me into these green covered SSRI [Social Science Research Institute] working papers—Goldberger's. Zvi was assembling sibling data sets. That was kind of a new thing. [The] longitudinal aspect was also kind of coming along.

I was lucky that I was there. There were new data, new data types showing up, and questions. But probably the main thing from Zvi was just a sense that he was in it because he was interested in the answers. He was open to any kind of methodological discussion and he was a very smart man. But he was driven by wanting to know things and some of that rubbed off, not enough. I think it makes a big difference whether you have some interest in knowing stuff as well as the methods. His openness to methods, well Savage was a contemporary—he was at Chicago while Zvi was at Chicago and Zvi would sit in on some of Savage's courses. I think there was one very abstract approach, to maybe the analysis of variances.[14] Basically Zvi said I do not care what the topic is, you just want to go and listen to this guy. Then I did hear Savage, partly there was some Bayesian conference activity that I guess Ed Leamer had organized. So, in Cambridge, Savage appeared. In terms of being influenced by the talk, yeah there was something very special here. I think the actual paper was probably what came out as proper scoring rules, the notion that you have these weather forecasters producing probabilities, like the probability of rain tomorrow. Well, what sort of compensation scheme should you set up so that it would be his dominant strategy to truthfully reveal what his probabilities really are. Of course, that had a history. There was a two-page paper, we should go check but perhaps McCarthy in Proceedings of the National Academy had the basic results and Savage had generalized it.[15] But yeah, that talk was—I remember that talk very, very well. So, I worked on the thesis under Zvi.

[12]Chamberlain, G. & Leamer E. (1976) Matrix weighted averages and posterior bounds. *Journal of the Royal Statistical Society Series* B, 73–84; and Chamberlain, G. & Leamer E. (1976) A Bayesian interpretation of pretesting. *Journal of the Royal Statistical Society* Series B, 85–94.

[13]Authors' note: Gary is referring to Leamer's 1978 book *Specification Searches: Ad Hoc Inference with Nonexperimental Data*, Wiley.

[14]Authors' note: see Alan Kruger and Tim Taylor's 2000 interview of Griliches in the *Journal of Economic Perspectives* 14, 171–189 for more on this class by Leonard Savage.

[15]Authors' note: the two papers being referenced are McCarthy J. (1956) Measures of the value of information. *Proceedings of the National Academy of Science* 2, 654–655; and Savage, L.J. (1971) Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association* 66, 783–801.

GI: Did Zvi have much interest in Bayesian things? With Chris and Ed there, there must have been a strong influence.

GC: I think that Zvi basically said that Bayesian answers were going to be the right ones, but he was put off by Zellner who he would have overlapped with at Chicago. And whether that was really all that conducive to empirical work. He basically wanted to get some empirical work done. That was not a big part of the interaction with Zvi.

BG: Was there an econometrics seminar at Harvard at that time?

GC: Yeah, it's always been Thursday, 4.30 pm to 6.00 pm. Zvi and Dale were running it, so yes. Or at least it quite possibly, most likely started with Zvi and Dale.

GI: So, if you compare it to the seminar now, how does the amount of theory and empirical work compare?

GC: Well, I guess no surprise it was much more empirically oriented. Can't say there are many of those [nowadays]. As a graduate student going to the joint Harvard-MIT econometric seminar, I do not have a lot of memories. Frank Fisher was very active, he was always there. OK so I remember one seminar, I do not recall who the speaker was, but Zvi had raised some point, and Frank said, now you have told him about this and now I will tell him that there was World War II, he should have taken that into account in his analysis. So, there was some back and forth. I do not have a lot of memories of particular seminars.

GI: In the early 1970s, did McFadden's stuff come up much? Chuck obviously was very influenced at that time. Was the discrete choice stuff as big a topic as panel data?

GC: OK so Chuck would have been working on his thesis under Frank Fisher at MIT. I remember Chuck saying that Frank was in some ways a great advisor. He would give him a draft and he would get back comments the next day. McFadden visited MIT during Chuck's period and was clearly very influential on Chuck.[16]

Another player would be Bob Hall: he was close to Zvi and he was at MIT. Not sure he ever published it, but Bob was interested in sibling related things. So, OK you have got observations from sib 1 and sib 2, and to what extent could you extract some latent variable that is ability related? He ended up actually looking at identification questions. I think I may have given him a reference—there is this book by Rao, Kagan and Linnik on random variables with linear structure.[17] They would have that sort of result, that in a linear factor model if you do not have normality, off the linearity and independence you can get all kinds of identification. Bob had a paper drawing on that, which may have had, well almost surely had, an empirical part. It would be kind of interesting, what the application was. Something like for birth there is a probability for a boy. Does that vary across families? He had data on multiple births, and [was] trying to use this linear factor

[16] See Tamer, E. (2019) The ET interview: Professor Charles Manski. *Econometric Theory* 35, 233–294.

[17] Kagan, A.M., Linnik Y.V., & Rao, C.R. (1973) Characterization Problems in Mathematical Statistics. John Wiley & Sons.

structure stuff, maybe the non-normal part, but probably all sorts of the simple covariance–variance stuff to get an estimate of how much heterogeneity is there in the probability of having a boy across families. I recall he had an estimate. This might have been joint with Dennis Carlton.[18] As I recall, they thought there was a significant but small difference.

BG: What about classmates from that time? Were there other students interested in econometrics?

GC: There were visitors. Laffont was visiting, possibly related to Zvi. But within the actual Harvard class, I do not recall a lot of econometric interactions. There might have been some graduate students from somewhere else doing work with somebody. Jacques Mairesse passed through. I am remembering Laffont passing through and a couple of other people that I will probably remember later. But no, there was not a close classmate. The classmate interactions might have been more getting problems to work on from some empirical problem. Jim Medoff would have been definitely a source of that. He was an intensely empirical labor person and would generate questions all the time. He was also a bit of a character.

GI: What did you write your thesis on?

GC: Latent variables, with applications to sibling data, also longitudinal data.[19] There was a joint paper with Zvi using some of the sibling data that he had assembled from the NLS, the brother's sample.[20] Some of it eventually kind of showed up in the panel data publications.

Well, OK there were things… there were these conferences. On the sibling stuff, Taubman and Behrman, they had twins from military datasets and there were various conferences organized around—well eventually there was a book that they edited—*Kinometrics*—and I think I ended up with some stuff out of the thesis in there with two or three pieces in there, partly identification stuff and sibling data and more generally on latent variable models.[21] Triangular systems with latent variables, and identification stuff; IV estimation, shortcuts to see where the identification is coming from. So, I guess that mainly came out of the *Kinometrics* volume that [Behrman and] Taubman edited. And these conferences were my first exposure to Goldberger. He would come to some of the conferences. So, these were the sort of conferences that were heading toward a book or a collection of essays. So, I knew Art from these green-covered SSRI discussion and working papers, but I got to actually experience Art [at these conferences]. And his critiques of some of the Taubman and Behrman work which were quite cutting, as in it appeared

[18] Author's note: the empirical application differs slightly from Gary's recollection, but we believe he is referring to Carlton, D.W. & Hall, R.E. (1978) The distribution of permanent income. In Z. Griliches, W. Krelle, H.-J. Krupp, & O. Kyn (eds.), *Income Distribution and Economic Inequality*. Halsted Press. In the article, the authors thank Gary for pointing them to Rao's theorem in Kagan, Linnik, and Rao (1973).

[19] Chamberlain, G. (1975) Unobservables in Econometric Models. Ph.D. dissertation, Harvard University.

[20] Chamberlain, G. & Griliches, Z. (1977) More on brothers. In P. Taubman (ed.), *Kinometrics: The Determinants of Socio-economic Success within and between Families.* North-Holland.

[21] Taubman, P. (ed.) (1977) *Kinometrics: The Determinants of Socio-economic Success within and between Families*. North-Holland.

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

that the authors had not read their own paper. Jim Heckman showed up at some of these; that was my first exposure to him [as well].

KH: What was the economics job market like back then for people coming out of the PhD program?

GC: I think I did not interview at very many places. At Johns Hopkins, I met Carl Christ. I think the main place I considered was Princeton, whether to go there. I decided to stay at Harvard and I just remember interviewing at a few places.

GI: Did you actually have to give a job talk?

GC: I gave a Princeton job talk, yes.

GI: Did you give a talk at Harvard as well?

GC: I gave a talk at Harvard. A certain flavor maybe would be captured—I think Ken Arrow was asked why Harvard hired some of their own students and he said, well we looked around and it seemed that we had some good people right here and we hired them. That would not happen today.

GI: Did anybody else join at the same time?

GC: Well, roughly the same time, but maybe a little later, Rachel McCulloch, who I later married. That worked out very well. And Janet Yellen was there, and she and Rachel collaborated and produced at least two publications together.

GI: Were they the first female faculty at [Harvard] Econ? Were there any senior female faculty at that time?

GC: No, but Leontief had his input-output empire. There were some women associated with it. Anne Carter had a long career at Brandeis, but she may have been an assistant professor at Harvard; I am not sure.[22] That would have been before I was a graduate student.

GI: But Janet and Rachel must have been among the very first?

GC: I would say so, certainly the ones that I remember, yes. Well, Graciela Chichilnisky was also at Harvard, but maybe visiting connected to work Arrow was doing.[23]

GI: What was the sense of being junior faculty there? Was there a sense that there was a reasonable chance of getting tenure at that time?

GC: No. There was some discussion [about my case]. Rachel and I got married and there was a two-body problem. [Moving to] Wisconsin, I think maybe it was Rachel who was the initial inspiration. Bob Baldwin, an international trade person, was very keen on hiring Rachel. But anyway, at some point, they started to gear up to see if they could make two tenured offers. That is when we started thinking about that. So, the wisdom of Zvi… There was some [interest from] Chicago, there was a question of maybe visiting Chicago for some period, whether that might have led to something. I think Zvi quoted me a probability of maybe 0.2 if I were to stay. He had heard some people at Chicago might be interested, maybe 0.4. He basically

[22] Anne Carter joined Leontief's Harvard Economic Research Project (HERP) in 1949. In 1969, she became the first woman on the faculty of the Economics Department at Harvard. See Akhabbar, A. (2011) Anne P. Carter: A biographical note. *Œconomia—History/Methodology/Philosophy* 1, 9–18. DOI:https://doi.org/10.4000/oeconomia.1805

[23] Graciela Chichilnisky was a Postdoctoral Fellow at Harvard in 1977, and a Lecturer from 1977 to 1978, before taking a position as Associate Professor at Columbia.

thought that Wisconsin seemed like a good deal and we eventually decided that we agreed. Our other main option then was UCLA. Ed [Leamer] was certainly an attraction, Finis Welch was at UCLA then, that was a major attraction. The chair Harold Demsetz was very aggressive at recruiting. That was not an attraction, the conversations with Harold. He had this—well his argument against Wisconsin was that to do good economics you have to be in a major city, you cannot do good economics in the sticks of Madison, Wisconsin. And Penn, we had joint offers at Penn, but decided to go to Wisconsin.

KH: So how did your thinking about econometrics and your research evolve in your time as an assistant professor at Harvard? Is this when you started doing the work that you are known for in panel data?

GC: A good bit of that kind of started with the thesis, and a lot of it very much influenced by Zvi. Then, so I guess my thesis was actually done in 1975. Then I was at Harvard as an assistant professor until 1979, something like that. Well, OK, so—I guess it was maybe shortly after I moved to Wisconsin, I need to check, but, Zvi was, of course, the editor of that handbook with Intriligator, and he had solicited the paper on panel data.[24] I would say a lot of the questions I kind of took with me to Wisconsin when I moved in 79. I probably got more interested just in the straight methods part of things at Wisconsin. John Geweke was there when I arrived. That was a very positive thing. But we lost him—two-body problem—he moved with his wife somewhere else. But Chuck Manski came and Jim Powell not much after that, and also John Rust, Ariel Pakes. The combination of, particularly I suppose, Manski and Powell. We will need to get back to what you were asking about McFadden and qualitative choice in that period, but I think with Manski and Powell in the environment, I got much more focused on econometric theory, just as methods in some sense for their own sake, which was not the kind of push I had been getting from Zvi. And it kind of came together. I tried to bring them together to some extent in the handbook chapter.

BG: There were a lot of different ideas in the panel data paper. One theme that is in the handbook chapter and the '82 paper,[25] is the idea that strict enough exogeneity generates over-identifying restrictions and using minimum distance as a way to combine these different restrictions. Is there a sort of point, an origin to that idea? Does it go all the way back to the thesis or does it grow over time?

GC: Well, I guess step one is this correlated random effects setup. There you definitely want to point to Yair Mundlak, who was visiting at Harvard while I was a graduate student and we interacted a lot. I think he was very close to Zvi and had an office in that part of the building. So that notion that you might try to build a link between the so-called fixed effects and random effects [was] very clearly in Yair's '78 *Econometrica* paper.[26] I guess part of what stimulated me was his conclusion

[24]Chamberlain, G. (1984) Panel data. In Z. Griliches & M. Intriligator (eds.), *Handbook of Econometrics*. North-Holland, pp. 1247–1318.

[25]Chamberlain, G. (1982) Multivariate regression models for panel data. *Journal of Econometrics* 18, 5–46.

[26]Mundlak, Y. (1978) On the pooling of time series and cross section data. *Econometrica* 46, 69–85.

FIGURE 2. Gary Chamberlain at Wisconsin, circa 1980.

that if it was properly done there was no issue. I was thinking about that, so I suppose, in particular, the notion of, well, what if there is a lagged dependent variable and you try to do the same things. Well, there is an issue. So incidental parameter problem type things, one way or the other, I got interested in through that. The autoregressive lagged dependent variable model—we were all influenced by [Stephen] Nickell. He had a piece showing the bias and inconsistency of least squares with autoregressions with so-called fixed effects.[27] But OK, so there is this correlated random effects notion coming in. That it would be useful to kind of have a multivariate regression view of things. I would say, Goldberger, or Goldberger and Jöreskog, that was coming from them.[28] The way they would set up latent variable models, with the MIMIC—multiple indicator multiple cause—this kind of multivariate way of looking at things in the panel context. Maybe it is siblings or maybe it is longitudinal, but basically you are saying that you have got those units, and you have got several variables attached to a person over time or to a family across the sibs. And treating that as a multivariate regression or some other flavor of multivariate analysis. I would say Goldberger-Jöreskog. And on that point, I very much remember Art Goldberger. I think it was when he introduced Tom Rothenberg as a seminar speaker at Madison once. He wanted to very clearly credit Tom as bringing econometrics into a multivariate view of modern statistics, in the sense of using multivariate analysis. On the identification side of things, there is also Vince Geraci, who was a Goldberger student at Wisconsin, and he wrote a thesis on identification issues, simultaneous equations with measurement error.[29] That was influential, I had a copy of that, I remember that quite clearly.

[27] Nickell, S. (1981) Biases in dynamic models with fixed effects. *Econometrica* 49, 1417–1426.

[28] Jöreskog, K.G. & Goldberger, A.S. (1975) Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association* 70, 631–639.

[29] See Geraci, V.G. (1976) Identification of simultaneous equation models with measurement error. *Journal of Econometrics* 4, 263–283.

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

GI: Wisconsin must have been an interesting place, with so many econometricians around, even if they did not all overlap. I think John Geweke had left before Jim Powell and John Rust and Ariel Pakes arrived.

GC: Yeah, I overlapped with John for maybe one year. Dennis Aigner had been there; I think he left just before I arrived. But yeah, there was a tradition somehow. I would connect it a lot with Art. The esteem and respect that basically everyone in that department had for him was quite amazing. I remember Mike Rothschild, who I overlapped with and we had some papers together, while he was in Madison. He had more experience than I had because he had been around more. He said he had never experienced anything like that, one person being such a focal point of a department. What was it? Having Art—it is kind of complicated, just what it was (Figure 2).

KH: So, you mentioned Mike Rothschild. Was that around that time you started getting interested in finance?

GC: Yeah, that was a one hundred percent due to Mike. He would ask me questions. So, I guess it started with getting the mean-variance—what sort of distributions or preferences would result in a mean-variance ranking of portfolios. Yeah, it was just basically a question from Mike, I would go off and think about it. So that had probably not a very important result but had a short, nice answer.[30] Probably more important—he was very influenced by Steve Ross in general, but, in particular, Steve's way of viewing that you could get in some ways similar results that you associate with CAPM from a factor structure-arbitrage approach.[31] And then lots of conversations about that, [to] understand the Ross results. And then quite a bit of—it was a period when that was the main thing I focused on. Totally due to Mike.

GI: Back to the Wisconsin culture there—what were the seminars like there? With Art Goldberger, Chuck Manski, John Rust, Ariel Pakes, Jim Powell, and you it must have been very interesting.

GC: Tended to be, I think there were department seminars, and we would not necessarily have a weekly metrics seminar, but sometimes the department seminar would be an econometrics speaker, and I guess sometimes we would have our own seminars. Art, he can be very witty, some of the things he said I remember. Day to day discussion was not necessarily driven by the last seminar speaker. There were other things, it might be what was Chuck working on or someone else was working on. I do remember a speaker, Art asked a question, and the speaker said, oh was that too fast? And Art said, it is not the speed so much, it is the direction.[32]

GI: That sounds like Art.

---

[30] Chamberlain, G. (1983) A characterization of the distributions that imply mean-variance utility functions. *Journal of Economic Theory* 29, 185–201.

[31] See Chamberlain, G. & Rothschild, M. (1983) Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica* 51, 1281–1304; and Chamberlain, G. (1983) Funds, factors, and diversification in arbitrage pricing models. *Econometrica* 51, 1305–1323.

[32] See also Kiefer, N.M. (1989) The ET interview: Arthur S. Goldberger. *Econometric Theory* 5, 133–160.

BG: So, was it at Wisconsin that you became interested in efficiency questions? Some of that is in the handbook chapter too.

GC: Very modestly. People in the department were interested in semiparametrics and I got interested. [I think] the environment had a lot to do with that.

KH: It seems like earlier Bayesian analysis was maybe more mainstream within econometrics and then around this time there is a lot more emphasis on frequentist econometric theory. Was that the case, and do you have a sense for why there was kind of a shift of a lot of people doing work on a more frequentist kind of mold?

GC: What were you thinking about in terms of Bayesian?

KH: Well, I guess in the 60s and early 70s there were a lot of people, Ed Leamer, you, Tom Rothenberg, there was a lot of work on Bayesian methods and then there seemed to be a lot of less of that in the late 1970s and 1980s. Maybe I am wrong about that but I am just wondering if you felt like that was the case and why that might have been?

GC: Well, certainly I remember the interest in semiparametrics, like Jim Powell's censoring work, Chuck's maximum score. Steve Cosslett certainly, we were aware of the work he was doing. That was very much frequentist asymptotics. Now in fact, it has a Bayesian asymptotic version, too. But no, I think it was the way one thought about semiparametrics that did not have any particularly Bayesian aspects. But I am not sure I had a sense of a lot of Bayesian activity, say my thesis was 1975, so around that time. There was Ed certainly, Arnold Zellner certainly I was aware of. But I do not actually remember a lot of Bayesian work at that time.

[break]

BG: I had some questions about the efficiency bound research. There are a lot of different connected ideas that I imagine might have been related to conversations you had in Wisconsin. First of all, the two papers—I am thinking of the 1986 and 1987 papers in the *Journal of Econometrics*,[33] there are two very different approaches to calculating the bounds and sort of where that came from? Particularly the latter approach using multinomial approximation, I would be interested in that, where this came from?

GC: Which was the first approach for me. That was really kind of straightforward. Suppose everything is discrete, not that many cells, so it is a very parametric problem. Fisher information should be just fine. So just combine that with methods of moments or with GMM, and there is a nice result there. The GMM part, that was in the air [with] Lars Hansen, Hal White, certainly by the time I moved to Wisconsin, maybe before. There was just a general awareness, that this would be a good way to set things up. So that, like I said, that was in the environment. And you get this nice result when you combine that with a multinomial, and that was me. I kind of feel good about that one.

[33] Chamberlain, G. (1986) Asymptotic efficiency in semiparametric models with censoring. *Journal of Econometrics* 32, 189–218; and Chamberlain, G. (1987) Asymptotic efficiency in estimation with conditional moment restrictions. *Journal of Econometrics* 34, 305–334.

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

GI: When you say it is straightforward, do you mean that it is straightforward once you set it up in the multinomial way? I always thought that the idea to set it up in a multinomial way was a very creative step, because in the end this really anticipated the later empirical likelihood stuff.

GC: Well, it is probably one of the papers I am most pleased about. I remember the first person I communicated the result to was Chuck Manski. We agreed to not noise it around until it got written up. The other way of doing things, well that was more when I started to get into the literature. This notion of a path, and so on. So, there I was just taking that and applying it to the censoring case and applying that to some models of econometric interest.

BG: The zero information result for maximum score answer, with the benefit of hindsight, many things seem obvious, but at the time was that surprising? Or was there a sense for which that was confirming an intuition that you thought… ?

GC: The business about the binary panel, two-period case, that [you have] identification essentially only if logit, so that was way back, maybe even Harvard, but certainly early Wisconsin. Actually, maybe at Harvard. So I knew that. Chuck's identification [of course] was clear under the conditions he was making, that you needed the infinite support.[34] Yeah that suggested that there could be an issue like that. But I guess looking hard to see what the information is in that case, well the motivation was Chuck. I was very aware of the maximum score result from Chuck.

GI: You just made a comment that the efficiency bound paper was one of the ones you are most proud of. Which other papers make that list?

GC: I suppose the lack of identification, the need for logit for identification.[35] There are things in the Fisher-Schulz lecture that I feel good about. I suppose the Handbook chapter. The work we did related to the Bayesian bootstrap[36]—well gee, that is the way I teach the first-year course, I mean a good chunk of the first-year course, I show other ways to get similar looking, get similar inferences. I think that it should have an impact. I am very interested—a new colleague Neil Shephard has interesting work on that sort of thing but you want to build in restrictions or combine that with a real prior,[37] so there is still a lot activity there. OK, the things that I like about the Fisher-Schultz Lecture… there is basically one idea that is maybe overworked at this point, but it would start with the many-IV-paper that we did[38] and at some point, though not initially when we were talking, but I guess it at least shows up as a sentence or two in the publication, the notion that there is this

[34] Manski., C. (1987) Semiparametric analysis of random effects linear models from binary panel data. *Econometrica* 55, 357–362.

[35] Chamberlain, G. (2010) Binary response models for panel data: Identification and information. *Econometrica* 78, 159–168.

[36] Chamberlain, G. & Imbens, G.W. (2003) Nonparametric applications of Bayesian inference. *Journal of Business & Economic Statistics* 21, 12–18.

[37] See Bornn, L., Shephard, N., & Solgi, R. (2019) Moment conditions and Bayesian nonparametrics. *Journal of the Royal Statistical Society, Series B* 81, 5–43.

[38] Chamberlain, G. & Imbens, G.W. (2004) Random effects estimators with many instrumental variables. *Econometrica* 72, 295–306.

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

invariant prior in the background that is maybe doing some of the nice things. Well, that was very much foreground in the Fisher-Schulz paper and then a paper with Marcelo Moreira[39] which just goes back to traditional panel data stuff under strict exogeneity in terms of trying to make a link between fixed and random effects. Well, that is the way I think the link ought to be made, that there is this totally fixed effects setup where you ask for some optimality, like in minimax or best invariant estimator, and just in the construction of that optimal fixed effects thing an integral appears. You do some integration and that can be reinterpreted as a particular random effect specification. Now in the paper with Moreira, there are different ways you can set it up, but one way looks very much like a traditional correlated random effects model. I kind of liked the reference I had there to a piece by Chris Sims,[40] which he gave I think it was at this conference in honor of Goldberger. And he had maybe three different pieces in his paper, but one piece was panel data, a kind of random effects autoregression and basically setting it up as a correlated random effects model, which was exactly what we could produce as an optimal fixed effects estimator under the invariance arguments. And so, yeah that idea I have gotten some mileage out of.

GI: Do you want to tell us something about how you moved back to Harvard?

GC: Well, I got an offer, and things worked out. Rachel thought we should try to go and then [she got an offer from] Brandeis. She thought there were pluses and minuses relative to Wisconsin for her, but the smaller classes were definitely a plus compared to some of the large state school issues. Initially, I thought it was a no-brainer, but at some point, it became a source of intense anxiety, that maybe this was a big mistake.

GI: After you were there or while you were making the decision?

GC: In the spring, I think probably after the interview with Mike Spence. I remember calling Rachel and saying I am not sure if I want to do this; Wisconsin has been a very good place for us. So actually, at that point it was perhaps Rachel digging in and saying, do it.

GI: What was it about Harvard that was worrying you? You had been there before, so you knew the place.

GC: It was more what I was giving up. Chuck was saying that you have got to at least—you do not have to resign, and that was understood, so at least try it, he said you have to try it, and that was sort of Rachel's view too. Now trying it involves buying a house, settling the kids in school... that is not so easy to reverse (Figure 3).

GI: So there at Harvard, were you teaching 2140b from the beginning?

GC: OK, so Zvi was teaching, there was a 2140 maybe ABC, a little vague. Basically, I split a course with Zvi and I split a course with Dale. That was all

[39]Chamberlain, G. & Moreira, M. (2009) Decision theory applied to a linear panel data model. *Econometrica* 77, 107–133.

[40]Sims, C. (2000) Using a likelihood perspective to sharpen econometric discourse: Three examples. *Journal of Econometrics* 95, 443–462.

FIGURE 3. Gary Chamberlain with Joshua Angrist in the Littauer computer lab at Harvard, circa 1990.

my teaching initially. That was interesting. Zvi had some wonderful qualities. Conscientious preparation of lectures was not one of them. So, there were issues there. Dale—there was no real effort I think to try to have a coordinated course. The course with Zvi, it was a full semester and we kind of alternated it, or maybe teaching with Ed Leamer.

With Dale it was a very clear division, so it was just like two half courses. I mean it seemed like a very light teaching load, half of one semester course, and half another. But at some point, I thought I would like to just teach a full semester of something. There was a bit of tension, and Zvi was pretty much always [saying] maybe next year we will do that, but at some point, we did it. So, teaching 2140 was a major part of Harvard for me for several years. Major major. Maybe a combination of the teaching fellow and a group of students who elected to take things in econometrics. But a really outstanding set of teaching fellows. I may confuse them a bit with students who played a large role in the class discussion but maybe was not a teaching fellow, but Kei stands out, Tom Knox, Aviv Nevo, I had you [to Bryan], yes [laughter], Parag [Pathak]. It really seemed like the ideal course. When I turned it over to you [Guido Imbens], I thought this was a gift.

GI: It was much appreciated. I really enjoyed…

GC: But, well, I also would say, some combination of my interests and whatever, but enrollment was if anything declining and not increasing, and the fact that when you took over, and enrollment really increased, that suggested this was a good move (Figure 4).

GI: It was a great course to teach. And similarly, I had a great experience with the TFs (teaching fellows) for that class.

GC: I had to turn Raj Chetty down. He wanted to be a TF, but there were already commitments. That was probably in his first year.

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

FIGURE 4. Gary Chamberlain and Guido Imbens in Monterey Peninsula, CA, 2002.

KH: So, your approach to teaching that course did change a bit over time?

GC: Yeah, and I think that was not helpful to the enrollment: more decision theory, while I think the median student was looking for more applications. The mix was going to the wrong direction. Teaching the first-year course is probably good, because it is sort of understood that [I am] trying to show you useful stuff, and I do not think there is a lot of extraneous material, at least from my point of view. Whereas in the way I was teaching 2140, I think taste would enter quite a bit more, in terms of, well, should you do this versus that. And I think my tastes were becoming somewhat less aligned with the median student (Figure 5).

[GI leaves]

BG: So, you touched a little bit about it in the context of talking about 2140, but did you have anything you wanted to say just about sort of advising and teaching, just more generally. Sort of how it's…

GC: I remember conversations with Rachel. Well, OK so after we had actually—we were here, the kids were settled, we were not going to go back to Wisconsin. So, now maybe go in a year or two years, it was very clear to her that a major benefit of me being at Harvard was the students. They were so

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

Figure 5. At Bryan Graham's dissertation defense in 2005 at Harvard University. From left to right: Caroline Hoxby, Gary Chamberlain, Bryan Graham, Michael Kremer, Lawrence Katz.

engaged and interested. I would say at Wisconsin I was more focused on my own research. And no question, interacting with Moshe [Buchinsky], Jin Hahn, Kei, you, Nick Barberis, was quite a major part of this research life. At Wisconsin, George Jakubson, we had quite a good interaction on some union panel data stuff with measurement error, and that was quite fruitful.

BG: And Kei also alluded to, in his question on the evolving structure of 2140, and also sort of in your own research, a more explicit decision theoretic framework. I guess it is always been lurking there, but sort of what prompted you to invest so heavily in that direction? Was it a philosophical sort of decision? Something you wanted to learn more about?

GC: There was a certain connection with micro theory that became somewhat more important to me. I am not quite sure why. Although I would say conversations with Al Roth stand out. And yeah, at some point I just really remember reading Blackwell-Girshick,[41] reading Luce and Raiffa,[42] and reading Wald.[43] A sense that this is a really good way to organize things. So, this, now I think, Chuck Manski is—certainly the Wald connection, was quite important to him. This would be post-Wisconsin; I do not recall us talking about decision theory much at Wisconsin. Later, when he would [come to] give a seminar we would talk quite a bit. OK,

[41] Blackwell, D. & Girshick, M.A. (1954) *Theory of Games and Statistical Decisions.* Wiley.

[42] Luce, D. & Raiffa, H. (1957) *Games and Decisions: Introduction and Critical Survey.* Wiley.

[43] Wald, A. (1950) *Statistical Decision Functions.* Wiley.

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

well, back to the Leamer influence, it was always back there. Savage's book[44] I have been through more than once, and it is just hugely influential, a huge influence for me. The connection with micro theory is part of this because getting exposed to Gilboa-Schmeidler would be a trigger.[45] That link between a piece of micro theory and minimax, and so on, just got me very interested. That you could have the Savage framework, you could drop one of the axioms and now you maxmin expected utility over a set of priors. So [at] this conference for Goldberger, the paper I gave was very much a Gilboa-Schmeidler maxmin [approach]. I think I probably stressed some computational issues, of finding the least favorable prior. I was very much into that at that time. I think part of it was just the notion of there is a piece of economics that is a natural link to statistics and that, gee, in econometrics you might try to exploit that. I think over time, I became a bit less attracted to kicking out that Savage axiom, because I would ask myself… I know about the Ellsberg paradox and I think it is probably a good axiom. And alternatives to single prior Savage decision theory. I think I am probably searching for other motivations. I think there is a role for that: something to do with the cost of carefully forming a subjective prior. I do not know how to formalize that very well.

BG: I remember when you gave the paper that eventually came out in the Oxford Handbook of Bayesian Econometrics.[46] The seminar participants sort of objected to your setup since you were claiming people are really making decisions this way, your response was that your decision maker had read Savage and was persuaded.

GC: Back to Al Roth, that, you know you would study normative versus descriptive. That you could be interested in how people actually make decisions. And you could try to study that in the lab and other contexts. I sort of thought what I was trying to do was to play the role of an advisor, that if some combination of patient and physician came to me saying how could you exploit these various kinds of clinical trial data, observational data, to make it a decision about this patient, well that is a real problem, that is a good problem so that I would like to do empirical work connected to that problem. That [paper in the] Oxford Handbook, a major weakness is it is not connected to a clear empirical dataset. And again, so back to the connections with economists doing decision theory, you need to keep clear, do you have a normative hat on or are you trying to say this is how the world works, and that is different.

BG: The normative social planner type motivation leads very naturally to Bayesian analysis and on one hand there is the problem of how do you encode prior information properly. And then, say Chuck's work, at least a big chunk of it sort of dispenses with the prior completely and leads to often less sharp conclusions.

---

[44] Savage, L.J. (1954) *The Foundations of Statistics.* Wiley.

[45] Gilboa, I. & Schmeidler, D. (1989) Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* 18, 141–153.

[46] Chamberlain, G. (2011) Bayesian aspects of treatment choice. In J. Geweke, G. Koop, and H. van Dijk (eds.), *The Oxford Handbook of Bayesian Econometrics*, pp. 11–39. Oxford University Press.

GC: I was very interested in Chuck's work—he had these bounds—he had a checkable condition under which you would basically do the finest level of matching.[47] That in terms of a treatment choice decision you would just look at people with the identical measured covariates, but then his student…

BG: You are thinking of Jörg Stoye.

GC: Yeah, a very nice result that in fact that was the minimax regret exact answer regardless.[48] I think the regret was added to minimax because sometimes minimax gives results that do not seem to agree with a plausible prior, but Chuck has taken the position that that is not fair, that the minimax regret stands by itself. But anyway, a motivation for adding regret to the minimax was, well, things like optimal decision rules being no data rules and yeah, it is interesting that that comes up with regret in that case, I find that very interesting. The bigger question is whether minimax is a useful way to approach things. I guess case by case, sometimes it seems to give useful answers. On the frequentist-Bayes decision theory intersections, there seem to be some people working on it. I would like to see [more of] that.

KH: I am curious here, it seems like there are a few application areas that you kind of come back to at various points in your work. One being returns to education, another being optimal consumption decisions and I am just wondering if you could just kind of talk about how you were thinking about these problems at different points. How have you approached those problems?

GC: The interest in returns to education, or more generally certain pieces of labor economics, that would go back to Zvi, very much so. Things like optimal consumption—well that just seems like a very basic thing for economists to be interested in. I guess my main memory there is frustration about the lack of data. People would work on food consumption with the PSID, and that is fine but it is kind of limited. That just seems like a very basic problem for economists to be interested in. I would like to be working on it with data.

KH: You have recently been working on some valued added, teacher valued added applications?

GC: In some ways I think about some of the quantile regression stuff as being stimulated by Larry Katz, he was doing work with Murphy and just in general there was a lot of interest in the department, maybe centered around Larry, on changes in returns to education, changes in inequality within education groups. And it motivated me to try to at least be pretty familiar with that literature and think about how to add to it. That led to work with Moshe [Buchinsky] and some papers that I like.[49] I was very pleased—Roger Koenker wrote me a letter after he saw the early version of the quantile regression paper indicating that he liked it, and that

[47] Manski, C.F. (2004) Statistical treatment rules for heterogeneous populations. *Econometrica* 72, 1221–1246.

[48] Stoye, J. (2009) Minimax regret treatment choice with finite samples. *Journal of Econometrics* 151, 70–81.

[49] See Buchinsky, M. (1994) Changes in the U.S. wage structure 1963–1987: Application of quantile regression. *Econometrica* 62, 405–458; and Chamberlain, G. (1994) Quantile regression, censoring, and the structure of wages. In C. Sims (ed.), *Advances in Econometrics: Sixth World Congress*, vol. I, pp. 171–209. Cambridge University Press.

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

meant a lot to me. And other big influences by colleagues, such as Raj's work on teacher value-added. I guess it partly started with some pretty specific questions. Well, OK I remember a party and Josh Angrist was there. And Josh was critical of an early version some of Raj's stuff. What was it? Allowing for some clustering in terms of a random effect specification, the standard errors actually Josh thought went in the wrong direction. So, I guess they got smaller instead of bigger. So, OK that is a little problem, trying to figure out was that a mistake or could that happen. We figured out that was okay, but then we just got stuck talking more about the sort of data he was getting his hands on. I have not published anything on the Project Star stuff, but we talked about that and well that data are, of course, quite public and that got me using the data a bit and thinking about things related to that data.

Teacher valued-added is a combination of getting interested through a colleague and some set of questions, plus this colleague having just unbelievable data that was not publicly available. And that was also a source of frustration, because I used the data and I published something based on it,[50] but I did not like the feeling of not being able to say here is the data, you see what you can do with it. The work Raj is doing with various people, but particularly the paper[s] with Nathan Hendren[51]—well I know because I had seen a bit of what goes into getting some of these data sets together, they made an effort to put very useful statistics that would preserve confidentiality on a project website. And I am using that [data] and I feel very good about being able to say go to that website and you can do what you want with that data. And again that would be exactly the same sort of thing: [my] colleagues' interest in spatial differences in mobility as a substantively interesting question. I attribute a lot of that to Raj, but here is a case where Raj and coauthors were actually just giving people what I think was a fabulous resource in terms of data. I think what makes econometrics interesting is if there is some interesting data maybe that is a slightly different type. When I was doing my thesis, panel data sibling longitudinal was sort of new. Neil Shephard with his second-by-second, tick-by-tick financial data, it just gives him lots of things to think about. I think that is what generates the interest in econometrics.

BG: Looking over the course of your career, different areas of econometrics have become more or less prominent over time. Are there any areas that you have been interested in but never had a question that you have been able to work on or things that you enjoyed following, things you wish you could have worked on but did not or wish you could have learned more about?

GC: When Al Roth was around—I am still not sure how important I think lab experimental economics is, but I sure felt Al Roth was super interesting. I acquired a fair amount of background knowledge in the experimental literature and tried a

[50] Chamberlain, G. (2013) Predictive effects of teachers and schools on test scores, college attendance, and earnings. *Proceedings of the National Academy of Sciences* 110, 17176–17182.

[51] See e.g., Chetty, R. & Hendren, N. (2018) The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *Quarterly Journal of Economics* 113, 1107–1162; and Chetty, R. & Hendren, N. (2018) The impacts of neighborhoods on intergenerational mobility II: County-level estimates. *Quarterly Journal of Economics* 113, 1163–1228.

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

few things in terms of what might have been a methods-oriented paper. Nothing has come out of that at least yet. Partly Rachel would make fun of me—why do not you work on some real interesting data, like spatial mobility.

KH: Changing the topic completely, you were a co-editor of *Econometrica*, you served on a number of other editorial boards and committees, for example for the National Academy of Sciences. Do you have any particular experiences from those that might be interesting?

GC: I guess I thought *Econometrica* was super important. The pluses and minuses: I would be amazed at the quality of some of the referee reports and the insights; minus I was not very good at delegating, so I just ended up not doing much else and at some point that did not feel so good. But yeah, I think the journal is super important. Committees, well certainly the NSF panel to award grants—that was a lot of work, but it was clearly a useful payoff. Very good group, very good discussions, so on the whole, I am very positive about that. I have to say committees in general, maybe it is just not—I certainly see some people that really seem to be in a committee meeting and they are somehow doing useful things, they are pulling things together. I have not had that sense myself much.

BG: We are getting towards the end. We could ask you a future of econometrics question. You may not want to answer it.

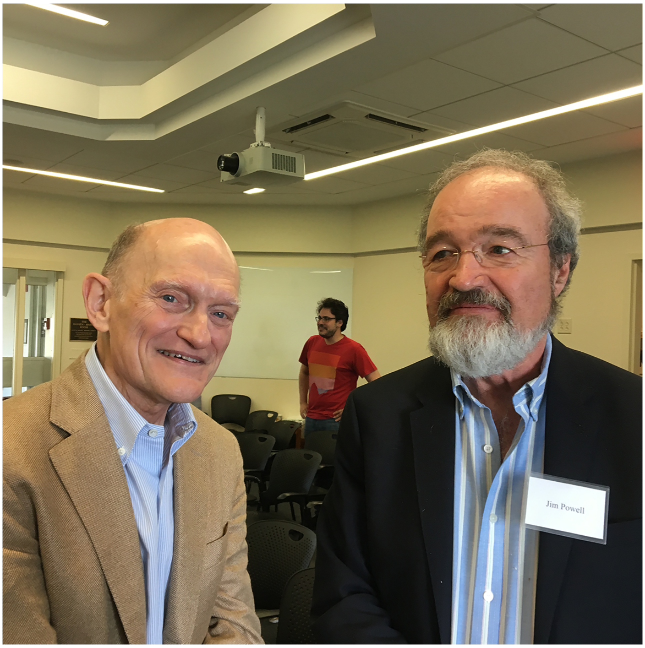


FIGURE 6. Gary Chamberlain and James Powell at the Conference in Honor of Gary Chamberlain, 2018.

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

GC: Apart from that it probably hinges on new, interesting new data sets, I do not have much to say.

KH: What are you most excited about in terms of your own research right now? Where do you see your work going in the next few years?

GC: I suppose trying to find some new interesting datasets. I have quite enjoyed working on this spatial mobility stuff and probably will do some more work. But it is very much, there is an interesting dataset. It is not just that, I mean I do seem to connect more with questions that are vaguely related to labor and parts of public, just as things that interest me, so there is that. Spatial things in general, that could be something that could grow out of the very particular spatial thing I have been looking at.

I am very interested in terms of the more methods-theory part, when there is a conflict between Bayesian inference and frequentist inference, even asymptotically. One of the things I remember most clearly from Arthur Dempster's course on multivariate analysis is the fact that first order parametric likelihood asymptotics, involving Fisher information, have this dual in the Bernstein-von Mises approximation to posterior distributions, though I do not actually recall Dempster using the term Bernstein-von Mises. Dempster felt that this was the most

FIGURE 7. At the Conference in Honor of Gary Chamberlain, 2018. From left to right: Ariel Pakes, Robin Lumsdaine, Parag Pathak, Keisuke Hirano, Moshe Buchinsky, Gary Chamberlain, Michal Kolesár, Bryan Graham, Thomas Knox, Jinyong Hahn, and José Luis Montiel Olea. Photo by Guido Imbens.

https://doi.org/10.1017/S0266466621000372 Published online by Cambridge University Press

important result in, maybe he said mathematical statistics, which was a term you would see then more than you see now. But yeah, I remember that very clearly. I think I was probably a sophomore, maybe a junior in college. But OK, in less parametric contexts, when there is some conflict where you do not have some kind of dual like that, I find that just intellectually super interesting. Well, OK: unit roots. I have no particular reason to work on unit roots; the empirical questions are not ones that I am close to. But the Sims-Uhlig paper,[52] and that sort of conflict, that is what I have in mind, in terms of here is a rigorous asymptotic posterior distribution where priors get dominated, where a whole range of priors that would lead to the same limiting posterior and that then might have credible sets that [look] quite different from what would seem to be the usual 0.95 asymptotic confidence set. But, anyway, Sims and Uhlig, that is the sort of conflict they were pointing to. And yeah, that is something I would very much want to follow if not contribute to, cases where there are clear conflicts like that. I would like to try to make it intuitive and try to have an argument for which is better if there is a conflict. Some of the recent work by Ulrich Müller and Andriy Norets[53]—it is in that general area, I find it very interesting. I suppose in terms of something I would like to know more about: essentially nonparametric versions of Bernstein and von Mises, and what are all the issues around that (Figure 6 and 7).

[52] Sims, C. A. & Uhlig, H. (1991) Understanding unit rooters: A helicopter tour. *Econometrica* 59, 1591–1599.

[53] See Müller, U.K. & Norets, A. (2016) Coverage inducing priors in nonstandard inference problems. *Journal of the American Statistical Association* 111, 1233–1241; and Müller, U.K. & Norets, A. (2016) Credibility of confidence sets in nonstandard econometric problems. *Econometrica* 84, 2183–2213.