Statistical Science
2024, Vol. 39, No. 2, 357–373
https://doi.org/10.1214/23-STS906

A Conversation with Guido W. Imbens

Fabrizia Mealli and Julie Holland Mortimer

Abstract. Guido Wilhelmus Imbens is the Applied Econometrics Professor and Professor of Economics with a joint appointment at the Graduate School of Business and the Department of Economics at Stanford University. He has made fundamental contributions to econometric and statistical methods for drawing causal inferences in experimental and observational studies, and applications to a wide range of disciplines beyond economics, including psychology, education, policy, law, epidemiology, public health and other social and biomedical sciences. Together with his longtime collaborator, Joshua Angrist, Guido was awarded half the 2021 Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel for their methodological contributions to the analysis of causal relationships, with the other half going to David Card.

1. ABOUT GUIDO: INTRODUCTION

Guido was born in Geldorp (Netherlands) on September 3, 1963, to Annie Imbens-Fransen and Gerard Imbens. When he was six, he moved to nearby Eindhoven where his father worked at the Philips electronics company, and later to Deurne where he went to high school. He has one brother and one sister. He attended Erasmus University in Rotterdam from 1981–1984, earned a Master's degree in Economics and Econometrics at the University of Hull in 1986 and a PhD in Economics at Brown University in 1991.

Guido taught economics at Harvard University (1990–1997, 2006–2012), UCLA (1997–2001) and UC Berkeley (2002–2006), before moving to Stanford in 2012. He is a Fellow of the American Academy of Arts and Sciences, the Econometric Society, the Koninklijke Hollandsche Maatschappij der Wetenschappen, the Royal Netherlands Academy of Sciences, and the American Statistical Association and a member of the National Academy of Sciences; he has received Honorary Doctorates from the University of St. Gallen (Switzerland) and Brown University. He is currently editor of *Econometrica*, the leading journal in econometrics.

Guido has advised or coadvised over thirty Ph.D. students and published more than 100 peer-reviewed articles. His book with Don Rubin (Imbens and Rubin, 2015) is an authoritative account of statistical inference about causal effects. According to Google Scholar, as of July 2023, Guido's academic work had received over 93,000 citations.

Guido is married to Stanford economist, Susan Athey, and the couple has three children: Carleton, Andrew, and Sylvia.

This interview was conducted on March 3rd and 5th, 2022, at Julie's house in Belmont, Massachusetts, and completed at various times over the following year.

2. YOUTH, FAMILY AND UNDERGRADUATE EDUCATION

Fabri: I'd like to start by talking about your youth.

Guido: I grew up in the Netherlands; we moved around, but stayed close to Eindhoven in the southern part of the Netherlands. The headquarters of the Philips electronics company were there at the time. Eindhoven was essentially a company town. Philips sponsored many of the local institutions, including the concert hall, public parks and the professional soccer team (PSV Eindhoven), and paid for the college education of the children of their employees, including my siblings and myself.

Julie: Because your dad worked for them?

Guido: Yes, my dad worked there his entire career. My mother worked there as well before she had kids, but stopped working to care for us. My father had gone to college for one year, but then he started working. My mother hadn't gone to college after high school, though she went later, when my siblings and I were in high school. My dad went back to college after he retired: he had studied physics when he was younger, but when he went back he studied literature.

Fabrizia Mealli is Professor of Econometrics, European University Institute, Florence, Italy, and Professor of Statistics, University of Florence, Florence, Italy (e-mail: fabrizia.mealli@eui.eu). Julie Holland Mortimer is the Kenneth G. Elzinga Professor of Economics and the Law, University of Virginia, Charlottesville, Virginia 22904-4182, USA (e-mail: juliemortimer@virginia.edu).

Both my siblings and I went to academic track high schools in the Netherlands. In the Dutch educational system, high schools are tracked starting at age 11, with vocational and academic tracks at various levels.

In high school, I enjoyed mathematics and related subjects; then it came time to apply for college. In the Netherlands, students don't apply to college in general, but apply for a particular subject. I wanted to do something close to math, but not pure math, partly because my older brother was doing that already. My economics teacher suggested that I look at econometrics, which was a separate academic discipline in the Netherlands set up by Jan Tinbergen in the 1960s. Curiously, there's no statistics undergraduate major in the Netherlands. The econometrics major partly substitutes for that; it was a mix of statistics, operations research and mathematical economics.

Fabri: I suppose you had very little statistics in high school?

Guido: Yes, I had reasonably high-level math in high school, but no statistics of any kind. But in the first two years studying econometrics, we spent probably 25 to 30% of our time learning statistics.

Fabri: So, liking numbers runs in the family?

Guido: Yes, both of my siblings studied mathematics. We were all interested in logic-type puzzles.

Julie: Do you think being a middle child had much of an effect on your choices?

Guido: Possibly. I think it did make things a little easier given that neither of my parents had a college degree. My brother was one year ahead of me and had done really well in high school. He established that going to university was expected and natural. My brother agonized over what he wanted to study, eventually choosing physics and mathematics. I wanted to do something different and go to a different place, which made econometrics an appealing choice. In Holland, universities specialize in different disciplines. In Utrecht and Leiden, you can study mathematics and physics, but only a little bit of economics; in Rotterdam, you could not study physics, but it was arguably the leading place for econometrics. I didn't visit many universities to make up my mind. I remember going to Rotterdam with my father and liking it there and that was pretty much it.

Fabri: How did you like the program in Rotterdam?

Guido: The program in Rotterdam was great for learning technical skills, but it was not so great for learning economics. I have this recollection about a macroeconomics course where we studied Keynesian-type models with markets for goods, money and bonds. At the end of the semester, I realized that I had no idea what bonds were and when I asked around, nobody else in my class had any idea either! We were all doing fine on exams because we could solve the problems without understanding what they meant.

FIG. 1. (*Left to right*) *Guido Imbens*, *Hanneke Imbens*, *Annie Imbens-Fransen*, *and Gerard Imbens in Rotterdam in* 1983.

Julie: How many years were you there?

Guido: I was there for three years. I never actually finished the program or got my degree. Wilbert [van der Klaauw], a friend of mine who also ended up in the US, read about an exchange program for students in Rotterdam to go to England. After our third year, we both went to University of Hull, which was a much more fun place to be a student than Rotterdam. It turned out it was also very good for us academically, because there was an econometrician there, Tony Lancaster, who was happy to have two Dutch students who were interested in econometrics. He put a lot of effort into teaching and advising us.

The first year, Wilbert and I did Master's degrees; we took regular courses and some reading courses with Tony. And then Tony asked us if we wanted to stay around for another year to work as research assistants and we both decided to do that. It was during that year that Tony suggested we think about doing PhDs in the US. And the idea of going back [to Rotterdam] to finish an undergraduate program having a Master's degree already seemed funny. Tony had just accepted a position at Brown University, so that made it easier to get into the PhD program there.

Fabri: Before we talk about the experience with Tony, do you recall any episodes or specific people that inspired you in high school or even at a younger age?

Guido: When I was in high school, my economics teacher gave me a little book by the Dutch econometrician, Jan Tinbergen, and that was inspiring because Tinbergen was a big figure in Dutch policy circles. People seemed to trust Tinbergen and he was an impressive person (see Dekker, 2021). He worked for the United Nations and earlier for the League of Nations, all while doing impressive scholarly work for which he won the Nobel prize [The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel, 1969]. After the war, he had been involved in setting up government institutions, such as Statistics Netherlands (CBS Centraal Bureau voor de Statistiek) and the Netherlands Bureau for Economic

Policy Analysis (CPB), both of which had bipartisan respect and credibility. For example, before an election, the CPB would take all the political parties' manifestos and run models to forecast what would happen if the government implemented those policies. Those forecasts would be taken seriously by the political parties and the voters—maybe too seriously!

Fabri: It's interesting that Tinbergen was the first Dutch economist to receive the Nobel Prize, and you are the third now.

Guido: That's right, [Tjalling Charles] Koopmans was in the middle [1975]. I read that Tinbergen was also nominated for the Nobel Peace Prize for his work with the United Nations. I think he and Keynes may have been the only economists nominated for the Peace Prize. I did meet Tinbergen when I was a student in Rotterdam when he gave a public lecture there, which made a big impression on me at the time.

3. EARLY INFLUENCES IN ECONOMETRICS AND DOCTORAL THESIS

Fabri: Let's continue with your academic career and your work with Tony Lancaster. I think the first time I read your name was in the book on transition data by Tony where he thanks you for helping him (Lancaster, 1990).

Guido: Yes—that was during my year as a research assistant in Hull. At that time, I did a lot of programming on discrete mixture models with duration data using the EM algorithm (Dempster, Laird and Rubin, 1977), which worked well in principle, but was incredibly slow. Tony had time on a super computer in London and I would waste hours and hours.

Julie: Were you taught programming, or did you teach yourself?

Guido: Programming was one of the things we were taught well in Rotterdam. We used to have courses that were completely devoted to that.

Fabri: What language did you use?

Guido: FORTRAN.

Fabri: The work you were doing with Tony, also later at Brown as a PhD student, was not about causal inference. Was it closer to the statistics literature?

Guido: Yes it was. Lancaster had brought the duration literature into the econometrics profession. Before then, econometricians had little interest in duration models, but then people started getting interested in Cox regression models and especially in the presence of unobserved covariates in these models, and how that would bias estimates of duration dependence. Implicitly, that was related to the causal interpretation of these models, but that was never made explicit.

Later it did feel there was a big difference between that early work and my later work on causality. I remember a conversation with Don Rubin around 1993 about instrumental variables and causal questions we were collaborating on. At some point, he asked what other things I was working on and I told him about the choice-based sampling work from my thesis. He asked, "Do you think of these models as causal? Or what do you think these models mean?" I got the impression he didn't think there was much value in that work at all! So for a while, those two parts of my research were very separate.

Fabri: What was your PhD thesis about?

Guido: My job market paper was on choice-based sampling (Imbens, 1992). One of the funny stories about that paper is that Josh [Angrist] didn't like my job talk at Harvard and was opposed to me being hired there.

Fabri: Josh interviewed you at Harvard!

Guido: Yes. He arrived at Harvard one year before me.

Julie: Was he a first-year assistant professor at Harvard at the time? Those are always the toughest ones!

Guido: Yes. He was outspoken in those days, and of course he still is! And he thought the work I was doing in my thesis was boring; I don't disagree with that now, looking back. I did another part of my thesis that I thought was more interesting on duration models, reversing the role of duration and calendar time. I made the calendar time dependence flexible and proportional to the duration dependence, which itself was modeled parametrically. That seemed useful in social science applications where the hazard rate may change over time in complex ways, for example, with the business cycle, but the duration dependence may have a simpler structure.

Fabri: Yeah, I remember that paper (Imbens, 1994)!

Guido: I thought that was a nice idea, but I didn't push it much further.

Julie: So did Gary Chamberlain have more vision than Josh in terms of hiring you?

Guido: Well, I think they were desperate to hire someone in econometrics! Two of their junior econometricians left that year, and they wanted someone to teach their econometrics courses. In fact, they didn't hire me in the usual way by first interviewing me at the American Economics Association meetings. Instead, they directly flew me out for a job talk. Clearly, I got lucky!

Fabri: I'd be curious to know the different characters of your mentors. I met Tony Lancaster once in England and everybody was scared to have him sit in seminars, whereas Gary seemed to be gentle.

Guido: Both of them were comfortable at thinking on the spot: you would have conversations with them and they would stop talking and think. They would be silent for what seemed like a long time, and then come up with an insight. That did make seminars intimidating, because they didn't ask clarifying questions or questions to make the speaker feel comfortable. They would ask deep questions such that the speakers or others in the audience weren't sure where the line of questioning was going. Neither of them intended to be intimidating, but it certainly had that effect on people.

But as a colleague, Gary was gentle and supportive of junior faculty. I remember when I got to Harvard, I'd written something and I gave it to Gary. He thanked me, but then I didn't hear anything. A week later, he left it in my mailbox completely marked up, line by line, with detailed comments both on content and exposition. He was impressive that way, incredibly thoughtful and he had a big influence on both his students and his colleagues, especially his junior ones. He was a great role model for the way I later interacted with my own students and junior colleagues.

In the work with Josh, Gary again played a big role. In the early version of the local average treatment effect paper (Imbens and Angrist, 1994), we had a selection equation that corresponded to a latent index crossing a threshold that determined whether an individual would receive the treatment or not, very much in line with the econometric literature of that time, for example, the work by Heckman (Heckman, 1990). I remember showing that draft to both Don and to Gary, and Don's reaction was, "This looks really messy, and if it's right, there's a much simpler way to do it." Gary said, "You could use potential outcomes here." And that suggestion made the argument crisp and clear. We thanked him for the comment, but it was a big insight. It allowed us to make that argument in three lines rather than have weird algebra and strange assumptions. And so both Don and Gary were right.

4. CAUSAL INFERENCE WORK

Julie: Was there a specific moment when you realized that, in a heterogeneous treatment effect world, an instrumental variable would measure a local average treatment effect?

Guido: That was a problem Josh and I thought about for quite a while. We saw this tension between what Josh had done in the draft lottery example (Angrist, 1990), which seemed very credible to us, and the existing literature. In that paper, Josh used the draft lottery as an instrument to estimate the effect of serving in the military on earnings later in life, with the key assumption being that the only way the draft lottery number affected later earnings was through its effect on veteran status. The formal results by Heckman (Heckman, 1990) and Manski (Manski, 1990) implied that if there was heterogeneity in the treatment effects, you couldn't consistently estimate the average effect; it was not identified. We felt there was a tension there, because clearly the effects were not constant in the draft lottery example, nor in most other applications.

Fabri: But the empirical results felt credible!

Guido: Yes, but exactly what made them credible was not clear to us. If the credibility was because the effect was constant, this seemed surprising. But we couldn't get there when we tried to understand the problem in a traditional latent index set-up and it took us a while to be clear about the problem and then write it in a potential outcome framework. Then there was a specific moment, I remember actually walking over to the [Harvard] Science Center to meet Josh, and realizing we actually had solved that part and essentially the paper was done. It was a fairly clear moment that we got to the insight, walking to the Science Center and realizing that this was not just one more step, but the final step.

Fabri: So this was in the early 1990s?

Guido: Yes, I think this was the summer after my first year at Harvard in 1991. Josh was about to move to Hebrew University in Jerusalem. Josh and I were colleagues at Harvard only for a year. I visited him at Hebrew University and he came back to Cambridge a couple of times, but the year we were both at Harvard, we spent a lot of time talking about these problems. That was a very formative period for me.

We were both living in Harvard housing, and we spent Saturday mornings at the laundromat in the building, talking about work. The Nobel museum asks everybody for an item that was meaningful in their research, and I donated a container of laundry detergent in recognition of those mornings.

Fabri: I guess at the beginning it was not all that accepted or well received?

Guido: That's right. There were two sides to that. One is that it didn't bother me much then, because the people whose opinions I valued thought it was interesting. This included Gary and Don. I remember presenting it at the joint Harvard–MIT econometrics seminar and there was a bigger crowd than normal; people had heard this work was interesting, and perhaps controversial and relatively accessible. So, we got positive feedback. But on the other side, there was pushback from other groups. I remember presenting the same work at a conference in Wisconsin where people were more negative.

It was annoying in a more narrow way. I was worried that we wouldn't be able to get the work published. But there we got lucky. I remember the decision letter from the editor on the local average treatment effect paper, who wrote something along the lines of: "the referees didn't really know what to make of this and they weren't sure what to recommend. And I agree with them, but I do think it is potentially interesting." And in this case the editorial process improved it considerably. The editor made us shorten it, and as a result it was much crisper and cleaner. It's hard to handle that type of paper (Imbens and Angrist, 1994) as an editor, so I respect the fact that the editor took a chance on it.

FIG. 2. *Items donated in Nobel Museum by David Card, Joshua Angrist (computer tapes) and Guido Imbens (laundry container).*

Julie: I would like to dig down a little bit on the parts of the economics literature where this has been influential, and the parts where there has been pushback. How do you view these different ways of doing research? Is some of the pushback, maybe what you heard in Wisconsin, valid in the sense that researchers may be asking different types of research questions, or studying different types of environments? And does that drive some of the pushback?

Guido: I think that's right. There are clearly cases where the local average treatment effect is not a very interesting object. I don't think we were trying to deny that. Recently, when I was at Berkeley, Peter Bickel told me a story about David Blackwell saying that he [David] wasn't interested in doing research, he was interested in understanding. That resonated with me. That paper, more than most of my papers, was really just about understanding the problem. It wasn't taking a position on whether compliers were an interesting subpopulation. It clarified what you got out of an instrumental variables analysis. We understood Josh' draft lottery paper much better after having written the local average treatment effect paper.

Similarly, in the quarter-of-birth application (see Angrist and Krueger, 1999), Angrist and Krueger were interested in estimating the return to schooling, the causal effect of an additional year of schooling on the logarithm of earnings. You can estimate a model that says that returns to education are constant across years of education and across individuals, but nobody really believes that assumption. The instrumental variables methods we developed allows you to be clear that under weaker assumptions you can estimate some average of the return for some subpopulation. And understanding what subpopulations these averages correspond to is clarifying. And in many cases, we are interested in different populations or what would happen in the future, so there is a sense that we are always extrapolating if we are trying to be relevant for policy. But we should understand the exact nature of the extrapolation.

I saw our work very much as a neutral thing, trying to understand what the analysis taught us. I remember being told that we should write it as a criticism of the existing empirical work, that researchers using instrumental variable methods were focusing on the wrong object. But that view was very different from my interpretation of the result. I felt it was about clarifying what these analyses were doing.

Julie: More of a constructive result.

Guido: Yes, but maybe I was naive about the research at the time. I wasn't trying to have a big impact; I wanted to understand some of the problems.

Fabri: I guess this relates to other work that you started while in Cambridge, all the methods developed under unconfoundedness, for example, Hirano, Imbens and Ridder, 2003, Abadie and Imbens, 2006. To what extent did you benefit by having these figures around you, Don and Josh and Gary? And do you think the field was prepared to receive this changing paradigm?

Guido: It was a funny time in a sense, because I didn't appreciate that what we were doing was different from the existing econometric literature. After the first paper with Josh I connected with Don, whose office was in the Science Center, the building next to the Economics Department. We started talking regularly, and then he suggested we teach a course together on causal inference. I decided I would focus my energy on that, and try to get the students in the Economics Department interested in it. I didn't really think about the possibility that it would change the field. It just seemed more interesting to me than what I was doing before. This sharing of ideas led to the AIR paper (Angrist, Imbens and Rubin, 1996).

In econometrics, the work on causality was viewed as not that important at that time. even much later I remember reading a review of Josh's book (Angrist and Pischke, 2009) that said: "It's a novel treatment of that sub-sub-sub-area of applied econometrics." I think that reflected the general attitude at the time, and for quite a while after that. But even though it would have been nice to have other people interested in the questions I was interested in, especially more junior people, it was fun to work on these problems. As a result of there being few people interested in this area, there was room for doing new things,

FIG. 3. (*Left to right*) *Guido Imbens*, *Don Rubin*, *Joshua Angrist in March* 2014.

because you didn't have to worry about ten other people working on the same thing. If you look at the causal inference area now, there's so much work going on, so many conferences dedicated entirely to causal inference every year, it is impossible to stay on top of it all. Any time you do something new, you worry that there may be other people somewhere else, possibly in a different discipline, working on the same questions; there is much more pressure.

Fabri: It must be exciting to look at that in retrospect.

Guido: Yes. In retrospect, I should have better appreciated the fact that it is great if few other people are working on the problems you're working on—assuming it's something that is *ex post* important. I remember a student, Phil Johnson, asking me: 'Do you really think this causal stuff is going to be important?" I said: "Yeah, I think so. I'm willing to bet my career on it," which seemed a bit of a wild statement at the time.

Harvard was a hospitable environment because the junior faculty were left free to do what they thought was important. When I went to the chair of the Economics Department, and asked: "Can one of my required courses be with someone in the Statistics Department on this new topic?" he said, "Sure, do whatever you think is interesting."

Julie: So, do you think these ideas were going to reveal themselves in one way or another? If you and Josh hadn't done it, would somebody else have eventually made the connection?

Guido: Yes, *ex post*, both Manski and Heckman were very close. If they had engaged more with Rubin's work, they probably would have gotten there too. I don't want to dismiss what we did, but it's clear that someone else would have figured it out at some point. Part of the challenge was studying what the other disciplines were doing and taking that into account, and that is always a hard thing to do when other disciplines use a different language.

Fabri: There were in fact similar results out there, but they had not been formulated in a general way.

Guido: Yes, Bloom had the one-sided case (Bloom, 1984) and Robins' work (Robins, 1986) also had basically the same assumptions, but he focused on the overall average effect, and got Manski-type results on bounds for the average effect. So, the step of getting to an average for a somewhat unusually *ex post* defined subpopulation required you to think of the question in a slightly different way. It's harder to do it, if you just have a parametric model.

Julie: How would you say the econometrics and statistics approaches have differed?

Guido: Don's early work (Rubin, 1974, Rubin, 1975, Rubin, 1978), which set up the problem in the form of potential outcomes, Y_t, Y_c and conceived of causal effects as the comparison of (i.e., the difference between) these objects was very important. That was a big contrast between the statistics literature as I saw it in Rubin's work and the econometrics literature at that time. That was ironic, because the econometrics literature was implicitly much more focused on causal effects from the beginning; you can see that clearly in the work by Tinbergen from the 1920s, and Haavelmo in the 1940s. But it approached the problems from a different way and that did not connect well with the statistics literature. But the small part of the statistics literature that was explicitly focused on causality established a very useful framework and came up with some key results. That also relates to the common way of doing empirical work in economics in the 1980s, where researchers often reported estimates of all the parameters they estimated and viewed all of them as if they were equally interesting. Now, the empirical literature has moved away from that. For example, in the empirical Industrial Organization literature, people are explicit in their focus on outcomes under counterfactual policies such as new regulations or mergers, and the specific parameter estimates don't mean much in these complex models.

5. STRUCTURAL TRADITIONS AND DIRECTED ACYCLICAL GRAPHS

Fabri: That brings up more structural traditions. If you think about Pearl's perspective, all the arrows from X to Y that define causal effects of several variables are less focused on a particular intervention (Pearl, 2009).

Guido: Yes, that's right. Often in economics, the starting point is one direct question: What is the effect of changing X on Y? The challenge is that isolating that effect may be difficult, but it is good to keep in mind that one often does not actually care about the other components of the model. In some sense, the DAG [Directed Acyclical Graph] approach is quite close to the structural tradition in econometrics going back to the 1960s. When I was a student in Rotterdam, we had one of these MONIAC [Monetary National Income Analogue Computer] machines, a big physical model of the economy built by William Phillips. Phillips, who was originally trained as an engineer, later became known for his work in macroeconomics on the Phillips curve (Phillips, 1958) that describes the policy trade-off between inflation and unemployment. The MONIAC machines had all these containers and pipes and levers, and you could simulate the economy by pumping water through the whole thing. In Rotterdam, they would demonstrate it to the students once a year. It was great fun watching it, because it generally wasn't working very well.

Fabri: Was it leaking?

Guido: Yes, no matter what you did, very quickly all the water would run out and the economy would crash. It was a great metaphor for how well these theoretical models worked. A lot of the theoretical models economists were building were similarly very elaborate, with lots of equations and accounting identities.

Julie: Lots of places to leak!

Guido: Indeed! At some level, these structural econometric models had the flavor of the modern DAG models, with lots of interconnected parts. Like the DAGs, the good thing was that the models made the links very explicit. But, in the 1980s, people became increasingly concerned that the credibility of these models was not very high.

Two papers captured these concerns eloquently. Edward Leamer wrote his famous paper "Let's Take the Con Out of Econometrics" (Leamer, 1983), arguing that nobody really believed empirical work in economics and that, while the models might look impressive, they were not credible. Their empirical results were sensitive to minor changes in the modeling assumptions, and he called for more sensitivity analyses. Later, when I was at UCLA, Ed Leamer was my colleague and we taught a course around these ideas. Like the earlier course I taught with Don Rubin, teaching with Ed Leamer was a wonderful experience, both for me and for the students.

Around the same time as Leamer's paper, Robert Lalonde wrote another famous paper "Evaluating the econometric evaluations of training programs with experimental data" (LaLonde, 1986); not as witty a title as Leamer's paper, but perhaps even more convincing. Lalonde showed that nonexperimental econometric methods were not able to replicate experimental estimates. He illustrated this by taking experimental data, putting aside the experimental control group and using a nonexperimental comparison group from some public use data.

Later, two students of mine, Rajeev Dehejia and Sadek Wahba, revisited Lalonde's data; their conclusions (Dehejia and Wahba, 1999), based on more modern matching methods, were a little different. But Lalonde's point was well taken and together with Leamer's paper it was the starting point for what Josh later called the "credibility revolution" in economics (Angrist and Pischke, 2010). It made many researchers in economics more skeptical of complex models and also motivated Josh and me to look for methods that could lead to more credible estimates. The two papers by Leamer and Lalonde were influential and I still teach them in my courses.

I do think that DAGs make sense in illustrating the critical assumptions in, say, an instrumental variables setting; for example, demonstrating what the exclusion restriction really means. But I never found DAGs very helpful for getting answers to the questions I was interested in, like the identification question for the local average treatment effect. The DAG literature has led to lots of new insights. Judea Pearl has been kind enough to give guest lectures in my PhD classes, first at Harvard, and more recently just before the pandemic at Stanford, and that was a lot of fun. Judea is a great presenter and the students found it fascinating and engaging; however, there still hasn't been much empirical work in economics using the insights from that literature. That is a bit curious, especially as the graphical methods have a long history in econometrics. Judea Pearl traces them back to Wright (Wright, 1934), but you see the causal graphs also in Tinbergen's work (Tinbergen, 1940), which even used the term "causal" to describe the graphs, and later in work by the Harvard econometrician Griliches (Griliches and Mason, 1972). Yet despite the work by these very senior econometricians, the graphical methods have not caught on in economics. Methods based on potential outcome approaches in contrast gained a foothold much faster. I wrote about some of the possible reasons behind that in Imbens, 2020.

6. BAYESIAN INFERENCE

Fabri: I would like to talk about Bayesian inference. You've advocated the use of Bayesian inference in general, not just in causal inference. In your recent discussion on the use and misuse of p-values, you state that the Bayesian approach of reporting credibility intervals is more coherent (Imbens, 2021). Why do you think, despite the elegance of the Bayesian approach in settings like partial identification, nonparametric modeling and sensitivity analysis, there seems to be skepticism and resistance about using highly parameterized Bayesian models?

Guido: I don't know. I'm sympathetic to Bayesian approaches. All my mentors, Tony Lancaster, Gary Chamberlain, Don Rubin were, or became more, Bayesian over time. I do see, at least in some places, more acceptance of Bayesian inference. In one consulting case, I helped analyze a specific experiment carried out in the context of a large number of experiments. Initially, the focus was on making decisions based on each experiment in isolation.

FIG. 4. *Guido Imbens at the MONIAC machine in Rotterdam in 2017.*

But eventually the decision-makers were persuaded that we could exploit the information from a large number of experiments by using Bayesian or at least empirical Bayes methods, and doing some shrinkage. It was an effective context to persuade them to take Bayesian methods more seriously.

In econometrics, Bayesian methods have not had a big impact because traditionally people were focused on doing asymptotics. That seems a bit of a cultural thing. Econometricians wanted to establish theorems for consistency and asymptotic normality. For many machine-learning methods, you can't obtain asymptotic results. This is one of the reasons it took a while for these methods to make a big impact in the econometrics literature; it faced some of the same difficulties as Bayesian methods. In some sense, regularization in machine learning plays a similar role to a prior distribution in a Bayesian analysis. So, I think that if you get people to use machine learning methods, they may be more open to Bayesian approaches as well.

Fabri: Right. It seems to me that in economics and in econometrics, more than in statistics, you have methods that all of a sudden become fashionable and everybody uses them.

Guido: Econometrics is a relatively small field. As a result, there is more agreement on what are the relevant questions and it can be harder to get people to accept new questions or new ways of looking at things. Whereas statistics is very broad in terms of applications, especially if you include the related parts of computer science, and there is room for people doing different things. That's also true about the causal inference literature these days; it has gotten broad, and you see that in the Online Causal Inference Seminar that I started together with Dominik Rothenhaeusler and Guillaume Basse, where people present work from very different perspectives.[1]

In defense of the econometrics literature, I do feel people have brought in new methods, sometimes prompted by the empirical literature, to start answering questions in new ways. Consider Alberto Abadie's work on synthetic control methods (Abadie, Diamond and Hainmueller, 2010, Abadie, Diamond and Hainmueller, 2015). It has become popular quickly, with applications and new theoretical work, and that methodology has now spread to other disciplines, including computer science.

Fabri: Yes, it definitely has.

Guido: It even gets referenced in the popular press, including The Economist and The Guardian, and there is now a huge amount of this empirical work going on in tech companies.

We saw something similar with regression discontinuity designs (RDDs). They had been around since the 1960s, coming from the psychology literature (Thistlewaite and Campbell, 1960), but they became more visible in the early 2000s after some influential applications (Black, 1999, Van der Klaauw, 2002, Lee, Moretti and Butler, 2004). That led to a flurry of empirical and theoretical papers. RDD methods are widely applicable, and the econometrics literature helped get it there. In the end, you do need compelling empirical applications to convince empirical researchers to invest in new methods.

7. SENSITIVITY ANALYSIS

Julie: Let's talk about sensitivity analysis, and this notion that we have yet to agree on all of the statistics that might be of interest.

Guido: I feel I have a great title for a paper here: In Search of the Third Number, with the idea that we, by and large, agree that the first two things we report after doing an analysis are a point estimate and a standard error, but there is very little agreement of what to report next. But I don't know what to write after the title! Some of these questions arose in meetings with decision-makers in my consulting work. In preparation, the data scientists would have done some experiment or analysis, and they would report a point estimate of some object of interest (say an average treatment effect), and some measure of uncertainty (say the standard error). But the rest of the meeting would be much vaguer and less structured. The decision-makers clearly have the capacity to understand more about how credible and robust these estimates were. But there is not a clear template to convey how credible

[1] https://sites.google.com/view/ocis/

these first two numbers are and what additional uncertainty there is about the counterfactuals. There have been many different versions of sensitivity analyses proposed and, in some sense, they are all trying to show how different the results could have been under different models and assumptions.

Fabri: Maybe Bayesian thinking and reporting of results is the closest you can get there, which include model uncertainty, Bayesian model selection and model averaging.

Guido: Yes, but it's hard to make that precise and to organize that in a way that you can apply in a routine way, without subject-matter knowledge or without it being context-specific. It seems quite remarkable that we do actually agree, in a lot of applications, on what the first two numbers are that you would want to see before making a decision. But what comes after that is more difficult.

Julie: Do you think different people would be asking for different things even in the same context?

Guido: Yes, suppose you do a randomized experiment and you get the point estimate and the standard error. What else would you say? Some people may be interested in analyzing treatment effect heterogeneity, maybe because this would make them more comfortable generalizing or extrapolating to different settings. Others may be worried about whether results are sensitive to the assumptions made or the statistical method that was used. What number would you report for the presence of heterogeneity? Or for sensitivity? There is no standardized way for reporting these phenomena and no systematic protocol for how to choose the space of models and methods one wants to explore, even though that seems very important. In practice, it is still very context specific.

8. MOVING TO STANFORD

Julie: Did the proximity of so many tech companies to Stanford impact your research when you moved there in 2012?

Guido: Yes, I had not expected that, and it came about in an indirect way. While we were in Boston, Susan [Athey] had been working with Microsoft, but I had not really gotten interested in the type of problems that people at the tech companies were working on. But then just as I arrived at Stanford, Jas Sekhon had started a data science seminar at Berkeley and we decided to run it jointly. From the beginning, we wanted to get people from different disciplines, but also people from the tech companies. And having the companies so close by made it easier to get industry people to give talks and attend. In fact, they were all within biking distance, Google, Facebook, Apple. Much of the research I've done in recent years has been inspired by some of the questions I saw there. This includes a couple of my papers with Susan Athey and Raj Chetty (Kolesár et al., 2015) where we look at combining experimental and observational data, and the surrogate paper (Athey et al., 2019).

One motivation came from the fact that the tech companies were often doing short-term experiments but were interested in long-term outcomes, and they wanted to make sure that they were not being misled by short-term outcomes. This comes up in marketing experiments where it's tempting to look at clicks as an easily measurable short-term outcome, rather than long-term engagement with the product. Optimizing for clicks just leads to clickbait problems [sensationalized text or exaggerated graphics designed to artificially increase clicks]. Gupta et al., 2019 lists this as one of the major challenges in online experimentation. Raj Chetty was interested in the same problem, looking at the effect of early childhood interventions on long-term education and labor market outcomes, and not just on—

Julie: On Kindergarten test scores?

Guido: Exactly! But from a statistical perspective, it's exactly the same problem. More recently, my interactions with the tech companies have made me interested in a set of experimental design questions. In a standard setting, we start with a population and then divide it typically into two groups, a treatment and a control group. But for a lot of the tech companies, there are multiple populations. At Uber or Lyft, for example, should you do experiments by randomizing the drivers or the riders? At Airbnb, should you randomize customers or properties? The insight was that you could think of the experimental design problem as corresponding to choosing a distribution of assignments on the matrix of driver-rider pairs: you could have some drivers who are almost always in the treatment group and other drivers who are almost always in the control group, and some for whom assignment is more evenly split. This can help to separate out the direct effects from the spillover effects of treatments on other riders or other drivers (Bajari et al., 2023, Papadogeorgou, Mealli and Zigler, 2019).

Spillovers are intrinsic in marketplaces, which are all about strategic interactions where people respond to incentives. That's different from Fisher and Neyman's settings, where there are plots of land that don't care whether you put fertilizer A or fertilizer B there (Neyman, Iwaszkiewicz and Kołodziejczyk, 1935). In settings with strategic interactions, the economic aspects of the problem require that you put some structure on the spillovers. That makes these experimental design questions interesting, when they do not address purely statistical questions, but they interact in an important way with the economics. Collaborating with statisticians and computer scientists helps to bring different insights and ideas to these questions.

Often we use experiments solely for answering direct decision questions: should we implement this intervention or not? Alternatively, you can think of any single experiment as providing incremental knowledge. You may want to think about combining the information from experiments with models based on subject matter knowledge to fill in the gaps where the models contain less information. So, you can think of experiments as just buying you information—

Julie: To identify key parts of the model.

Guido: Exactly, and not just for one particular question, but to learn about the overall structure of the problem.

9. WORK WITH TECH COMPANIES, DATA ACCESS

Fabri: We would like to ask about your consulting work with tech companies.

Julie: And specifically data access. I'm sympathetic to the fact that the tech companies have a liability to keep data secure. But they have become a gatekeeper to data on much of the economy now, and this can make it difficult for independent academic researchers to have access to information that's important for understanding policy effects. Do you see a way of making sure that independent research doesn't get shut out?

Guido: Fairly early on when I was at Stanford, I spent a summer at Facebook. Formally, I was an intern. I had a desk and would go there a couple of days a week. I was working with Dean Eckles who is now at MIT, and Eytan Bakshy. The setup there was like many tech companies, just big open spaces. In the area where I was working, there were probably 100 to 200 people, all sitting or standing at their desks. It reminded me of pictures from early 20th century textile factories, where you see women standing at rows of sewing machines. All the desks were the same, probably 70% used them as standing desks and 30% as sitting desks, just row after row of people typing on their computers.

Fabri: I would hate working by standing!

Julie: You could be one of the people sitting down!

Guido: Yes! You could move your own desk up and down; that was an individual choice. But it does create differences compared to people sitting in separate offices. One is that people were not surfing the web; people were working. And it was nice to be able to ask questions to people nearby about what you were working on. At some level, the interruptions could be annoying, but it was also nice to have people immediately around. If people didn't want to be disturbed, they would wear headsets.

At that time, Facebook had done a couple of interesting experiments, such as looking at how a change to people's news feeds affect what they subsequently post. It's clear that Facebook needs to think about what to put in your news feed. They want to show things that are relevant to you and not show the same thing to everybody.

There was one result that if you show people happier pieces of news, they're more likely to post happy things themselves. Of course that seems likely to be true, and it's important to know what the implications are. There's psychology research showing that people are affected by others' behavior. Facebook got bad press about this at the time, which I didn't quite understand. But the bad press clearly made it attractive for Facebook to say, "We are not giving out this type of data, or publishing this type of research." And presumably, my guess is that, they're doing some of it now just internally.

Julie: Did they do the research at the time to write up formally into papers or just for internal presentations?

Guido: I think that research was done by a mix of people from the outside and from Facebook and published in a journal (Kramer, Guillory and Hancock, 2014). The outside researchers had strong incentives to publish the research, and Facebook must not have considered it a big problem at the time. But once a company becomes large it's easy for others to give that type of research a negative spin. At that point, company lawyers are likely to say, "This activity has a lot of risk and little upside."

It was not so different from when I worked with Don Rubin and Bruce Sacerdote (Imbens, Rubin and Sacerdote, 2001) on the lottery. When we first went to the lottery commission in Massachusetts, they had had some bad publicity and they thought our research would be a way of showing that people who win the lottery are happier and doing well. They still went to their lawyers to ask about risk, but then they gave us access and we surveyed people. At some point, though, one of the people we surveyed complained: "Why is the lottery giving out my address to these researchers who ask me invasive questions?" Then the lottery commission was reluctant to engage further. In all these cases, the risk for the companies is potentially large and the upside potential is very small.

Julie: From the organization's perspective. From the perspective of academic progress, the upside could be big.

Guido: That's correct. But the way the system is set up now, the companies might get a little bit of good publicity, but pretty much all of it is downside risk for them. There's no reason for someone to take that risk because it can only come back to haunt them. They're never going to get any credit for giving out the data.

Julie: In that telling of the story, it seems things have not changed much.

Guido: I think that early on, when companies are small, they don't have lawyers to worry about the downside risk—and the risk is small because the companies are small. But as soon as you have a regular legal department, someone's going to say, "There's no reason to do this." Early on, if you talk to the founding generation of these companies, they're always doing things that are high risk and they don't care that much.

Julie: You think the founders have a greater tolerance for risk?

Guido: Yes, I do. I think the benefits are more tangible for them. They're actually interested in learning about these things, and they value the expertise of outside researchers who can help them with their questions. But over time, I think the risks loom much larger than the benefits. Now some students do internships at these companies so they can use the data there, but getting data outside is very hard. The privacy regulations now, especially in Europe, make it much harder and make the burden of proof much higher.

Fabri: Although I do understand the risk for these firms, at the same time, there is also a social responsibility.

Guido: Yes. But those benefits are not internalized by the lawyers. The legal departments are directly worried about antitrust issues and reputational risks. The Facebook example is convincing. You do need to decide what to put in peoples' news feeds but you'll never stop people from saying that you're manipulating behavior. If you look at traditional newspapers, there is space devoted to things that make people feel good. They don't solely publish the most newsworthy things.

Julie: You're pointing out the fact that although the size of the companies may change the nature of the problem, these are longstanding issues. And on that point, there are things that can be done, like anonymizing data, that still allow researchers to make progress on important issues. Maybe anonymizing doesn't solve the concerns in some cases, but how optimistic are you that outside researchers will be able to get access?

Guido: A fundamental problem is that all these companies have objectives. The objectives may be good for society, but they may also have negative impacts on groups. Uber or Lyft have been bad for taxi companies, no question. And so research could reveal effects that make the company look bad or good. If you do an analysis of data from Uber or Lyft and you write a paper that shows how many taxi drivers they've put out of work, that's going to look bad. So, the companies have an incentive to allow research that emphasizes the good things and does not emphasize the bad things. There are probably many benefits from having platforms like Uber and Lyft, in terms of workplace flexibility or road safety. And so you can imagine that if you're Lyft, you would be happy to give access to data for people to study those issues but not for studies of the effect on taxi drivers. This is not even a privacy issue. The data could be completely anonymized, but there's still potential for a message about the company that creates awkward incentives.

Julie: Well, that's been the case even for data sets that are being made available to academics. The policy landscape shifts and all of a sudden, even if you had access to data, you're no longer allowed to study antitrust issues, for example.

Guido: I think that's really more than the privacy issue. It is hard to see how you would convince the companies or make it attractive for them to allow for that, because for the researchers there's an incentive to do things that generate public interest, and public interest may be heightened if there's some clear negative or positive effect.

10. ROLE OF MACHINE LEARNING: PREDICTION VS. CAUSAL INFERENCE AND CAUSAL DISCOVERY

Fabri: We already touched on this topic, but I wonder if you want to say more on the role of machine learning, especially in causal inference.

Guido: The two biggest changes I have seen in econometrics since I have been in the profession are causal inference and machine learning. Machine learning has improved many of the things we used to do poorly. We used to do nonparametric regression using kernel methods, and we knew that it was not working very well as soon as there were more than a couple of variables. For supervised learning (or prediction) problems, the new methods are vastly superior. This has direct implications for causal questions where prediction is also involved, because if we can separate the problem into causal and prediction components, we can import better prediction methods into causal inference.

But I think there is also a shift in perspective and emphasis. In econometrics, in particular, we used to be focused on methods for which there was a particular set of properties—consistency, asymptotic normality and no asymptotic bias—so that we could construct confidence intervals. This is what you see in journals, and the econometrics profession has probably overemphasized this. Machine learning methods have made it clear that we are not necessarily always interested in inference. For pure prediction- or classification-type problems, it may be sufficient to have good out-of-sample properties. I think this is an important insight that the econometrics literature paid little attention to previously.

Now, out-of-sample properties are much harder to establish for causal questions because fundamentally we cannot directly verify or validate them. In many cases, we do need to find ways of adapting these machine learning methods to causal settings and I think they are going to be incredibly useful there.

Other parts of the computer science literature do different things, like causal discovery methods. It's still hard for me to see exactly where causal discovery is taking us; asking why something happened, instead of what the effect of something is. These problems are called by Andrew Gelman "reverse causal questions" (Gelman and Imbens, 2013), and we often see them in policy and consulting settings. "Something went wrong; why did it go wrong? Was it because of X, or because of Y, or because of something else entirely?" We haven't studied them in the causal literature very well. One difficulty is understanding exactly what is meant by that question: how far back do you have to go to have a satisfactory root cause or satisfactory explanation?

Fabri: In some settings, you may have a better idea of how far back you have to go. When using genomic data, for example, it is a closed set of variables that you can consider. Causal discovery may not tell you exactly what the cause is, but it can be suggestive of possible causes.

Guido: Yes, there is room for progress to be made. The causal discovery results are probably going to be helpful in doing that, and this is a place where the interdisciplinary nature of the field is useful.

11. CAUSAL INFERENCE IN MACRO

Fabri: Let's discuss the role of causal inference in macroeconomics.

Guido: In macro, we typically don't have units such that we can think about an experiment or about varying the treatments, because there's a single economic system and an equilibrium, and if you perturb the system in particular ways, there are effects on multiple sets of agents, and we need to unpack those effects and incorporate whatever equilibrium we think is right.

In the 1990s, I remember Christie and David Romer (Romer and Romer, 1994) looked at the effect of monetary policy on output. Previously, people had looked at these questions by running regressions of, say, output on interest rates. The Romers used the minutes of the Federal Reserve (Fed) meetings and saw that there were times when it was clear the Fed was going to raise or lower interest rates. They argued that you couldn't learn much from those episodes because those events had already been taken into account in the economy. But if, on the day before a Fed meeting, the Wall Street Journal didn't know what the Fed was going to do, then you could look at the subsequent reaction to see what the causal effect was. That gives something closer to an experiment and you get closer to estimating a causal effect. I felt that work was very much in the spirit of causal inference in the microeconomic literature.

For instance, it relates to some of the microeconomic studies of the perception of race and gender on labor market outcomes. Claudia Goldin and Cecilia Rouse (Goldin and Rouse, 2000) studied the effect of gender discrimination in orchestra auditions by looking at auditions where the musicians were behind curtains, so the committee couldn't see what different musicians looked like. Marianne Bertrand and Sendhil Mullainathan studied racial discrimination (Bertrand and Mullainathan, 2004) by changing names on resumes in a way that would affect perceptions of job candidates' race, and then analyzed the causal effect of the name changes. In all these cases, thinking about something you can actually change, even if it is just the perceptions that you can change, and not the reality, can be helpful for understanding what you can learn from observational data.

12. THE ROLE OF PRIZES, AND CULTURAL DIFFERENCES BETWEEN ECONOMICS AND STATISTICS

Fabri: How do you think prizes affect the culture or environment in different fields?

Guido: Prizes are a good thing and a bad thing. It's clear that they draw a bright line between researchers that doesn't do justice to reality. There are often many people working in areas and even if you could agree on the merits of particular contributions, there is no reason to draw a bright line between one contribution and a slightly better or worse contribution. That is the awkward part of prizes. At the same time, prizes are clearly effective in bringing attention to fields. From that perspective, I think the Nobel Prizes in 2019 (for Abhijit Banerjee, Esther Duflo and Michael Kremer for experimental evaluations in development economics) and 2021 (partly for causal inference) brought attention in economics to causal methods; similarly with the Frontier of Knowledge and Turing awards that Judea Pearl won for causal inference, and the Rousseeuw prize that Jamie Robins, Eric Tchetgen-Tchetgen, Thomas Richardson, Miguel Hernán and Andrea Rotnitzky won, also for causal inference.

In general, since 1969 when they added the economics prize to the set of Nobel prizes, I think it's overall been very good for economics. It generates outside interest in what people are doing; economics has been good at managing prizes and getting publicity for them, such as the Clark medal for the most influential economist under 40. What the optimal number is; that's tricky. You want a small enough number that prizes get outside attention. In statistics, there isn't a single prize that generates the same kind of attention.

At one of the Nobel events, someone from the Nobel Foundation asked David Card and myself, if they were to ever add Nobel prizes in other areas, what would be good areas? Both David and I said that Data Science and Computer Science were arguably the most natural areas to add. I don't know whether they ever will; they haven't added any prize in 50 years or so. But prizes can generate attention and funding to basic research. And to the credit of the Nobel Foundation, they've managed the prize very well, with effective outreach to high schools and other places.

Fabri: On a more personal side, what are the privileges and responsibilities you face with the prize?

Guido: In the end, there are a lot of good things and not really any bad things that come out of it. The most challenging thing is that there is pressure to weigh in on topics that you don't know much about. I get asked to do

FIG. 5. *(Left to right) Carleton Imbens, Susan Athey, Guido Imbens, Sylvia Imbens, and Andrew Imbens, early morning after the Nobel Prize announcement, October* 11, 2021.

things that I didn't get asked to do before, but most of those are fun. It's easier to get in touch with people or organizations that I want to talk to. And you get to meet interesting people! I met the reggae artist Shaggy at the 2022 Brown University commencement and he was really fun to talk to.

Fabri: Do you feel pressure that, going forward, your work is going to be more scrutinized?

Guido: No. I enjoy doing research and I don't think it's going to be better or worse now. I remember talking to one of the other Nobel laureates in medicine, who had had a grant application rejected the week after the prize. In the 5 months after the prize, I had five papers back from journals: one very grumpy revise and resubmit, and four rejections. I don't think it is going to get any better, but I think it takes away the right to complain!

Julie: How do you think the cultures of economics and statistics differ?

Guido: The cultures of the two fields are quite different. Statistics is broad because so many fields have a connection to statistics. That's led to different types of departments. Sometimes, statistics departments teach all the statistics courses, including for all the separate fields, and sometimes they only teach their own students, and all the other departments teach their own statistics courses. That leads to a culture where it's not always clear what the role of subject-matter knowledge is and how big a role that plays in the methods people use. It also has advantages, with people coming in with different backgrounds and questions.

FIG. 6. *After the Nobel ceremony in Stockholm, (left to right) Andrew, Guido, Sylvia, Carleton, December* 10, 2022.

In econometrics, the connection to economics has been important from the beginning. People like Jan Tinbergen were focused on problems that naturally arose in economics: for example, how you separate out the demand function from the supply function. This is fundamentally not a question you can answer without thinking about where the equilibrium concept is coming from. Personally, I have always found it helpful to have the connection to economics to motivate some of the problems. I think for statisticians in general, it's helpful to have a connection to a substantive field, whether that's biology or a social science, to motivate and ground the questions.

13. WRITING, CHOOSING RESEARCH PROJECTS, EDITORIAL WORK, STUDENTS

Fabri: Do you have any memorable experiences of writing, either of papers, like Angrist, Imbens and Rubin, 1996, or of the book with Don (Imbens and Rubin, 2015)?

Guido: Two of my big influences in terms of writing were Don and Josh. They were both adamant about particular parts of their writing styles; where the commas went, what the sentence structure should be, what the paper structure should be. They would keep rewriting until they had things the way they wanted them. I learned from them to enjoy the process of writing and I try to get things to the point where I'm happy. That's not necessarily always successful! But I do value that part of the job. The early papers with Josh took a lot of time in terms of writing. And for the book with Don, we kept writing and rewriting it to make it feel right. We started talking about writing it in 1995–1996 when we were teaching our causal inference course. But we both kept getting distracted by other projects, so it took a long time to get to the finish line. But it was very enjoyable!

Julie: How do you manage your editorial work at *Econometrica*, and how do you choose which research projects to work on? I always find in my own work that choosing which projects to work on is easier than choosing which projects not to work on. So, there's this triage process every morning when you wake up. There's always 100 things to do, and you only have time to do 10 of them.

Guido: The editorial work takes a lot of time; it's a constant flow. It's hard when you travel and you don't get some of it done and there's immediately a backlog. But at the same time, it's important. It's the way the profession works. I feel by doing that, I can have an impact by publishing work that I think is good, and by making the work that we do publish better. Editors do not necessarily get credit for helping other people, but I don't really need that at this stage of my career. I have had more credit than I need, or possibly than I deserve, and I like the editorial work as a way of influencing the direction the profession takes.

And yes—choosing things to work on is always hard. Sometimes projects seem like they will be useful to work on and I get excited and put effort into them for a sustained period of time. And other things take years to figure out how to do and sometimes die a slow death because I lose interest in them. That's always hard, especially if it's with coauthors because they may have different incentives. It's hard to figure out when to stop a project, and it's hard not to get involved in too many things. I try to commit the night before to what I'm going to do the next day but I rarely manage to stick to it.

Julie: Does it stress you out?

Guido: No, I don't get stressed about it. It probably stressed me out more at other stages in my career, even though when I look back now, there were many fewer things to worry about, and fewer projects to choose between. Now, deadlines come and go.

Julie: Let me ask you about students, which is a funny thing for me to ask you about, I guess!

Guido: Which ones are my favorite students? That's like, "Which one is my favorite child?"

Julie: Exactly! At each of the different universities you've been at, you've connected with a new group of students. What impact have students had on your research?

Guido: Working with students is one of the fun parts of academic jobs. Teaching is fun, but it's more intense with students you're advising; seeing them grow, starting off nowhere close to the research frontier, then learning about that and then going beyond the frontier. Seeing that growth process is gratifying. The other thing that's fun is working with them and learning from them; seeing them come up with things that you didn't know. That's been true at all the places I've been. At each place, it took a couple of years to have a group of students who were interested in the areas I was interested in. But it would always end up being fun. That's probably the hardest part of moving on to other places, leaving students and having to start from scratch again.

14. WOMEN IN STATISTICS AND ECONOMICS

Fabri: Neither statistics nor economics has many women. But you've coauthored with women and mentored women. And you're married to a prominent female economist, Susan Athey. So you've had different front row seats over the years. How has the profession changed for women?

FIG. 7. (*Left to right*) *Orville Burrell* (*aka the reggae artist Shaggy*) *and Guido Imbens getting honorary degrees at Brown University commencement*, *May* 28, 2022.

Guido: I don't think we are there yet, but it has changed a huge amount. I remember when I was a junior faculty member at Harvard, I was at a seminar dinner and the chair of the department and his wife were there, as well as two female junior faculty and some others. The wife of the chair at some point started talking about how she thought women should not be working in the labor force, and they should be staying at home and looking after children. I was a very junior person, but the chair did not say anything supportive, given that there were two female junior faculty there, both trying to do their research and get tenure. And this now seems an astonishing situation and incredibly inappropriate from the chairman's perspective not to call that out. Relative to that, I think we have changed a huge amount.

But many changes are very recent. At Econometrica, we have six or seven coeditors serving concurrently. Until I became editor at *Econometrica*, there had only been one female coeditor in the 90-year history of the journal. Now we have three female coeditors and a fourth one coming on board soon. We also recently had the first female president of the society [Pinelopi Koujianou Goldberg, President in 2021], which had never happened in the previous 90 years of the society.

Graduate programs are now much more balanced in terms of gender, but it's not there yet in terms of senior faculty, where most departments still have very few women. And you see it in other parts of the profession—only two women have won the Nobel prize in economics, both fairly recently, compared to, say, eight in chemistry, some, like Marie Curie, going a long way back.

One of the things I see at *Econometrica* is that reviewers tend to be aggressive and negative in a way that's not supportive of junior researchers. It's not necessarily that they're more aggressive against female authors, but it creates an environment that I think many people, and many women, in particular, are not comfortable with. So, one of the things I'm trying to change at *Econometrica* is to make sure that reviews are more professional and less personal.

Julie: More constructive. It's hard to be constructive.

Guido: Well, that's true, but it's easy to be more polite. And I think that's one place to start.

15. MARRIAGE AND PARENTING

Fabri: I'd like to ask about your work with your wife. Susan (Athey) is also an active, high-profile researcher. How do the two of you support each other and juggle everything?

Guido: Well, our responsibilities at work are a little different. Most days, I have fewer things that cannot be changed last minute. Teaching responsibilities cannot be changed. But other than that, most of my meetings and days are filled doing editorial stuff, which cannot be postponed indefinitely, but can always be moved from one day to the next. And I don't like to cancel student meetings, but if I move a student meeting to the next day, that's not a disaster. I like to keep my days flexible enough so that I can move things around.

Before the pandemic, Susan tended to have a lot of travel, which generated inflexibility in her schedule. My schedule tended to be more flexible with less travel, so that helped. And we have a full-time nanny, even though the kids don't need that much help so that helps with logistics and last-minute things. It also helps that Susan and I have common research interests. Susan is broader than I am, but I understand the things she's working on. We discuss what we are both doing on a daily basis, so that helps create a supportive environment. Her work is more stressful than what I do. She has a lab, and managing that creates inflexibility and pressure to deliver.

Since the Nobel prize, some of the things I'm doing are less flexible and there's been more demand on my time, and Susan has been incredibly supportive. She knows what some of these things entail, such as giving large public lectures, so she's been helpful in giving advice.

Julie: Has parenting played a role in the way you approach mentoring or research? Or has it changed the way you engage in your work?

Guido: That's an interesting question. None of the kids are particularly interested in economics. They all have different interests. So, parenting has been good as a way of focusing on different things and being forced to take a break. Living in California makes it easy to go to the coast and hike on the beach or go kayaking. Parenting keeps you sane.

Julie: Someone once asked me what I do to get inspired to teach undergrads. I said, "At dinner, my kids often ask what I'm working on, and I'll explain my lecture to them. And if there's any part of that where I'm not connecting with them, they identify it immediately." There's a sense in which it does directly affect my work.

Guido: Yeah! For the Nobel prize, Stanford did a video where I explained to my kids what I was doing.[2] One of my colleagues is now using that in his classes; apparently, it helped to explain the content more clearly than I was doing before.

Julie: What do your kids make fun of you for?

Guido: I think they make fun of us for being so obsessive about our work. But I think they generally view us with some amusement. You see that in the YouTube video with the kids' interviews.

16. FINAL REMARKS

Fabri: What do you think is going to be important going forward in econometrics? Are you optimistic about certain directions of the field?

Guido: I think it's an exciting time to be doing econometrics, and data science and statistics in general. And I think what makes it exciting is partly the interdisciplinary nature of the work now; there are many substantive areas where there is interest in sophisticated data analysis. I think combining methods from different areas and collaborating across different disciplines is promising.

Computer scientists are doing a huge amount of interesting work in causality and causal discovery. I think, at some level, one of the biggest challenges is finding good ways of bringing in subject-matter knowledge in a way that improves machine learning and purely statistical methods and finding approaches that are subject-matter specific.

Julie: Do you see statistics and econometrics being more customer-focused, in that sense, going forward?

Guido: I think some of that is happening now. In econometrics certainly, there is more interaction between people doing empirical work and people doing pure econometrics. And that's been good.

ACKNOWLEDGMENTS

The authors thank Dan Ackerberg, Kei Hirano, Joseph Hotz, Phil Leslie, Alessandra Mattei, Aviv Nevo, Joris Pinkse, Geert Ridder and James Stock. They also thank Fiammetta Menchetti and Arnab Palit for editorial assistance.

[2] https://www.youtube.com/watch?v=uKlfqTz4-q4&t=34s

REFERENCES

ABADIE, A., DIAMOND, A. and HAINMUELLER, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *J. Amer. Statist. Assoc.* **105** 493–505. MR2759929 https://doi.org/10.1198/jasa.2009.ap08746

ABADIE, A., DIAMOND, A. and HAINMUELLER, J. (2015). Comparative politics and the synthetic control method. *Amer. J. Polit. Sci.* **59** 495–510.

ABADIE, A. and IMBENS, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica* **74** 235–267. MR2194325 https://doi.org/10.1111/j.1468-0262.2006.00655.x

ANGRIST, J. and PISCHKE, J.-S. (2009). *Mostly Harmless Econometrics*: *An Empiricist's Companion*. Princeton Univ. Press, USA.

ANGRIST, J. D. (1990). Lifetime earnings and the Vietnam era draft lottery: Evidence from social security administrative records. *Amer. Econ. Rev.* **80** 313–336.

ANGRIST, J. D., IMBENS, G. W. and RUBIN, D. B. (1996). Identification of causal effects using instrumental variables. *J. Amer. Statist. Assoc.* **91** 444–455.

ANGRIST, J. D. and KRUEGER, A. B. (1999). Empirical strategies in labor economics. In *Handbook of Labor Economics* **3** 1277–1366. Elsevier, Amsterdam.

ANGRIST, J. D. and PISCHKE, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *J. Econ. Perspect.* **24** 3–30.

ATHEY, S., CHETTY, R., IMBENS, G. W. and KANG, H. (2019). The surrogate index: Combining short-term proxies to estimate long-term treatment effects more rapidly and precisely. *Working Paper* **3895**.

BAJARI, P., BURDICK, B., IMBENS, G. W., MASOERO, L., MCQUEEN, J., RICHARDSON, T. S. and ROSEN, I. M. (2023). Experimental design in marketplaces. *Statist. Sci.* **38** 458–476. MR4630378 https://doi.org/10.1214/23-sts883

BERTRAND, M. and MULLAINATHAN, S. (2004). Are emily and greg more employable than lakisha and jamal? A field experiment on labor market discrimination. *Amer. Econ. Rev.* **94** 991–1013.

BLACK, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *Q. J. Econ.* **114** 577–599.

BLOOM, H. S. (1984). Accounting for no-shows in experimental evaluation designs. *Eval. Rev.* **8** 225–246.

DEHEJIA, R. H. and WAHBA, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *J. Amer. Statist. Assoc.* **94** 1053–1062.

DEKKER, E. (2021). *Jan Tinbergen* (1903–1994) *and the Rise of Economic Expertise*. Cambridge Univ. Press, Cambridge.

DEMPSTER, A. P., LAIRD, N. M. and RUBIN, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B* **39** 1–38. MR0501537

GELMAN, A. and IMBENS, G. (2013). Why ask why? Forward causal inference and reverse causal questions. *NBER Working Papers 19614, National Bureau of Economic Research, Inc.*

GOLDIN, C. and ROUSE, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *Amer. Econ. Rev.* **90** 715–741.

GRILICHES, Z. and MASON, W. M. (1972). Education, income, and ability. *J. Polit. Econ.* **80** S74–S103.

GUPTA, S., KOHAVI, R., TANG, D., XU, Y., ANDERSEN, R., BAKSHY, E., CARDIN, N., CHANDRAN, S., CHEN, N. et al. (2019). Top challenges from the first practical online controlled experiments summit. *ACM SIGKDD Explor. Newsl.* **21** 20–35.

HECKMAN, J. (1990). Varieties of selection bias. *Amer. Econ. Rev.* **80** 313–318.

HIRANO, K., IMBENS, G. W. and RIDDER, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica* **71** 1161–1189. MR1995826 https://doi.org/10.1111/1468-0262.00442

IMBENS, G. (2021). Statistical significance, p-values, and the reporting of uncertainty. *J. Econ. Perspect.* **35** 157–174. https://doi.org/10.1257/jep.35.3.157

IMBENS, G. W. (1992). An efficient method of moments estimator for discrete choice models with choice-based sampling. *Econometrica* **60** 1187–1214. MR1180239 https://doi.org/10.2307/2951544

IMBENS, G. W. (1994). Transition models in a non-stationary environment. *Rev. Econ. Stat.* 703–720.

IMBENS, G. W. (2020). Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics. *J. Econ. Lit.* **58** 1129–79.

IMBENS, G. W. and ANGRIST, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica* **62** 467–775.

IMBENS, G. W. and RUBIN, D. B. (2015). *Causal Inference—for Statistics*, *Social*, *and Biomedical Sciences*: *An Introduction*. Cambridge Univ. Press, New York. MR3309951 https://doi.org/10.1017/CBO9781139025751

IMBENS, G. W., RUBIN, D. B. and SACERDOTE, B. I. (2001). Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players. *Amer. Econ. Rev.* **91** 778–794.

KOLESÁR, M., CHETTY, R., FRIEDMAN, J., GLAESER, E. and IMBENS, G. W. (2015). Identification and inference with many invalid instruments. *J. Bus. Econom. Statist.* **33** 474–484. MR3416595 https://doi.org/10.1080/07350015.2014.978175

KRAMER, A. D., GUILLORY, J. E. and HANCOCK, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proc. Natl. Acad. Sci. USA* **111** 8788–8790.

LALONDE, R. J. (1986). Evaluating the econometric evaluations of training programs with experimental data. *Amer. Econ. Rev.* 604–620.

LANCASTER, T. (1990). *The Econometric Analysis of Transition Data*. *Econometric Society Monographs* **17**. Cambridge Univ. Press, Cambridge. MR1167199

LEAMER, E. E. (1983). Let's take the con out of econometrics. *Amer. Econ. Rev.* **73** 31–43.

LEE, D. S., MORETTI, E. and BUTLER, M. J. (2004). Do voters affect or elect policies? Evidence from the US house. *Q. J. Econ.* **119** 807–859.

MANSKI, C. (1990). Non-parametric bounds on treatment effects. *Am. Econ. Rev. Pap. Proc.* **80** 319–323.

NEYMAN, J., IWASZKIEWICZ, K. and KOŁODZIEJCZYK, S. (1935). Statistical problems in agricultural experimentation. *Suppl. J. R. Stat. Soc.* **2** 107–154.

PAPADOGEORGOU, G., MEALLI, F. and ZIGLER, C. M. (2019). Causal inference with interfering units for cluster and population level treatment allocation programs. *Biometrics* **75** 778–787. MR4012083 https://doi.org/10.1111/biom.13049

PEARL, J. (2009). *Causality*: *Models*, *Reasoning*, *and Inference*, 2nd ed. Cambridge Univ. Press, Cambridge. MR2548166 https://doi.org/10.1017/CBO9780511803161

PHILLIPS, A. W. (1958). The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861–1957. *Economica* **25** 283–299.

ROBINS, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math. Model.* **7** 1393–1512.

ROMER, C. D. and ROMER, D. H. (1994). Monetary policy matters. *J. Monet. Econ.* **34** 75–88.

RUBIN, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* **66** 688–701.

RUBIN, D. B. (1975). Bayesian inference for causality: The importance of randomization. In *Proceedings of the Social Statistics Section of the American Statistical Association* 233-239.

RUBIN, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Ann. Statist.* **6** 34–58. MR0472152

THISTLEWAITE, D. and CAMPBELL, D. (1960). Regression-discontinuity analysis: An alternative to the ex-post facto experiment. *J. Educ. Psychol.* **51** 309–317.

TINBERGEN, J. (1940). Econometric business cycle research. *Rev. Econ. Stud.* **7** 73–90.

VAN DER KLAAUW, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression–discontinuity approach. *Internat. Econom. Rev.* **43** 1249–1287.

WRIGHT, S. (1934). The method of path coefficients. *Ann. Math. Stat.* **5** 161–215.