[THEME MUSIC]

SARAH HANSEN: Today on *Chalk Radio:* how economics has become a powerful and collaborative tool with real-world impact.

ESTHER DUFLO: I realized that economics was the answer to my dream of having an academic career, while having an influence on the world right here and now.

SARAH HANSEN: For this episode, I'm joined by two illustrious guests-- Dr. Sara Ellison, a senior lecturer in the MIT Economics Department, and Professor Esther Duflo, a Nobel laureate and co-director of J-PAL the Jameel Abdul Latif Poverty Action Lab. Sara and Esther have spent much of their careers looking at how the field of economics can aid in providing equitable outcomes and better policies using data. Together, they developed the course 14.310x, Data Analysis for social Scientists, which you can find on MITx and MIT OpenCourseWare.

Sara and Esther were both in France at the time of this interview. But lucky for us, they were willing to connect and explain how they got into economics in the first place, why the field might be outgrowing the term "economic" and a new vision for the field that is highly collaborative. We also discuss AI and tackle the question, is it possible to visually represent data in a way that is both honest and compelling? We loved this conversation, and we hope you enjoy it as much as we did.

Welcome to both of you. I'm so interested in both of your economics origin story. I'm curious about how you got your start in the field, what sparked your interest, where that came from. Esther what about you? Where did this interest come from?

ESTHER DUFLO: So for me, it started a little bit later in life, in that I started doing my undergraduate degree in history, first in general social Science then in history. I did a master's degree in history. And during that master's degree, I was actually spending a whole year in Moscow to research my master's degree dissertation. And I had the opportunity of being a research assistant for some economists who were working in Russia.

I also had the opportunity to meet Thomas Piketty, who was at the time actually an assistant professor at MIT, but spending quite a bit of time in Russia. And through this experience and this meeting, I realized that economics was the answer to my dream of having an academic career while having an influence on the world right here and now, which history was only satisfying the first half of that plan. So that's when I decided to transition to economics and came back to do a master's, and then move to MIT for my PhD.

SARAH HANSEN: Interesting. Thank you. And Sara, what about you? Where did your interest in economics come from?

SARA ELLISON: Well, my back story is maybe the opposite of Esther's. I don't know. I can remember wanting to be an economist since I was in eighth grade.

SARAH HANSEN: Really?

SARA ELLISON: Which is extraordinary, because I didn't know any economists and I didn't really know what economics was. I just had this notion it was a way of thinking that was appealing to me. It was sort of logical, but not as rigid as maybe the way that mathematicians think. And that appealed to me a lot. And the fact that the frameworks and the constructs that economists use could be applied to social problems was really appealing to me. Now, I say that again, emphasizing, I really didn't know what economists did, and I really didn't know what I was getting myself into. But I'm stubborn enough that I stuck with it.

SARAH HANSEN: So I mean, you bring up a really good question. For people outside of the world of this podcast and in this room-- and actually including myself, most of the time, like, what is economics? If you're not in the field, what is it, and how does it help the world?

SARA ELLISON: Yeah, economics-- I suppose there are some textbook definitions out there. But the way that I think about it is, it's a set of models, it's a set of frameworks, it's a set of tools. And we use all of those models and frameworks and tools to try to answer questions about the way the world works and the way that people behave and the way that firms behave, and who comes out ahead and who comes out behind when things change.

That's how I think of economics now. It's really a very broad field. And maybe we're even outgrowing the term economics in some sense.

SARAH HANSEN: Esther I'd love to get your take on that. Do you think that the field is outgrowing that term?

ESTHER DUFLO: Well, one very short answer for what is economics is whatever economists do. And sometimes, other social scientists who think economists have imperialistic tendencies use that definition in a disparaging way. And I use this definition in a loving way, which is I think there are no problems that don't have some amount of economics in it.

So when you define it this way, then, to become just slightly more narrow, you can link it to the way of thinking that Sara mentioned. And I think maybe one way to differentiate it from sociology, for example, is that it's a lot of the same problems we are thinking about. But economists tend to think about what are the consequences of some specific things, where sociologists tend to think about what are the causes of specific phenomena.

SARAH HANSEN: Interesting.

ESTHER DUFLO: So in that sense, we are, in terms of mindset, almost closer to medicine or to the natural sciences in having this linear notion of you have a manipulation in the world. What effect will it have on the individuals who are directly affected by this manipulation, on the individuals that are close to them, or on an entire system?

SARAH HANSEN: Yeah, that's very interesting. And it made me think a lot of your research, Sara, in industry and industrial organizations. And you've done a lot of work in pharmaceuticals and e-commerce. What is the effect of the decisions that these industrial organizations are making on everyday people?

SARA ELLISON: So I guess one of the themes that runs through a lot of my research in industrial organization has been how do firms decide how to set prices. And that obviously has huge implications on consumers and on overall levels of welfare. And so I've studied this in various settings. I've studied it in pharmaceuticals, so asked and tried to answer questions about how does political pressure on pharmaceutical firms change their pricing behavior.

How does proximity to an upcoming patent expiration change pricing behavior? How does the market power of the retailer who's purchasing the drugs change pricing behavior? And then I've also done that in other settings, like e-commerce, asking the question of why are prices online not as low as lots of people expected them to be? A lot of people at the dawn of e-commerce expected that people would be able to go online and do price searches, and that would force prices really low and the internet would be a bargain basement. And that didn't happen.

SARAH HANSEN: It doesn't feel like that, no.

[LAUGHTER]

SARA ELLISON: It didn't happen. And so another set of questions I've studied is looking at, market by market, why didn't that happen, what are the forces going on there. So just to give you one quick example, there are several online markets where firms can use internet technologies to make prices harder to find and to thwart price search, and to obfuscate what their products and prices are.

And there are certain circumstances when they would have an incentive to do that. And that's obviously not good for consumers. So that gives you a little bit of a taste of the kinds of things that I've studied.

SARAH HANSEN: Yeah, that's so interesting. I mean, when I've considered economics in the past, I've thought of it largely as like a theoretical field.

And I really wasn't aware of the social justice aspects of it or the impacts on individuals and communities. This is really interesting. So thank you for opening my eyes to that.

And I know, Esther, you won the Nobel Prize for economics in 2019, along with your colleagues. And you were looking at poverty alleviation. And I'm wondering if you could tell us a little bit about your research, what that process was like, and what you found, and how it's impacting people now.

ESTHER DUFLO: Yes, so my work is definitely very empirical, as is Sara's for the most part. And it's about trying to understand different aspects of poverty. So often, when people think about global poverty, poverty in the world, the problems become so large that they either become discouraged and decide to stop thinking about it, or they think about radical solutions that are supposed to address the entire problem at once.

For example, I was just on radio just before this interview where they kept asking me, but don't you think the problem is with capitalism ultimately? Which it might be. But hey, I'm not so sure. And instead, my strong belief is that one needs to go problem by problem-- a problem we understand by problem we understand. Cut down this huge issue that is poverty into manageable issues, both as a matter of intellectual pursuits-- because once we have a manageable issue and well-posed questions, with a bit of luck, we can find the right answer, and also in terms of policies, because policymakers, all they have are their one-by-one policies in their toolkit.

So what I started doing pretty early in my career, and what J-PAL, the lab that I co-founded, has become known for is to set up randomized controlled trials to test the effectiveness of policy interventions. So a randomized control trial is an experiment very much the same way that you would test the effectiveness of a new drug or a new vaccine. There is a treatment group that receives the new vaccine and a control group that doesn't.

And they are randomly assigned, so you are sure that the difference between the treatment and the control-- if there is any, is due to the vaccine or to the treatment. So until several decades ago, it wasn't used much or at all for policy instruments, mostly only in agriculture, in fields to test various seeds against each other, or in the labs or for drugs. And what we've developed with J-PAL with Abhijit Banerjee with Michael Kremer and with many, many, many others, because that's a big movement, is the use of these experiments on social problems.

For example, if you want to know whether it would help to have computers in schools, then you could take 200 schools and randomly assign 100 to get the new computer and 100 to continue business as usual, follow them for a year, and know what has happened to the computers and to the teaching and to the kids at the end of a year. And you would have learned a lot on whether the computers are useful, which would give you an idea of whether this is something we could scale up. And also, I think little by little, you get a sense of-- for example, in this case of education, why the learning levels are low in the first place, and what are the set of interventions that are likely to be successful, which gives you the idea for the next things you might try.

SARAH HANSEN: So your work must be highly collaborative. I imagine in a study like that, you need to work with educators, learning scientists, anthropologists, people who understand how to measure those changes in the field that you're looking at.

ESTHER DUFLO: Yes, our work is deeply, deeply collaborative. In fact, the very day, we got the Nobel Prize we started saying, and haven't stopped since then, that it's really for a movement rather than for any individuals. In all honesty, I think if some alien came and zapped out every single paper with my name on it, it wouldn't really change anything of substance. Whereas I do believe that as a whole we've achieved something.

But any individual paper or project is less important than the mass of them. And it is absolutely true that every single project is a big collaboration in the same way that it is in natural science because so many people are involved. So if you take, for example, an education project-- we've started working with cognitive psychologists for the measurement of cognitive skills across different domains, both mathematics and language, et cetera, but also social skills, ability to understand the theory of minds of other people, and so on, including with very small kids.

For example, we can now do testing of children as young as 18 months, which is not something economists would be very good at designing, but which in collaboration with psychologists, we can do. We, of course, need to collaborate with people on the ground, which are either NGOs, but often also government workers, and in the case of education, teachers-- of course, the parents and the students themselves, a whole category of people who are collecting the data, who is also a profession, and so on and so forth. So project has a large number of co-authors. And that too masks even larger teams of people who are involved in the project. So that project becomes more of a little enterprise than someone sitting at their desk in front of a white page and trying to master deep thoughts.

SARA ELLISON: What Esther hasn't quite told you is that she is in the vanguard of young economists who have completely transformed the field of economics to this kind of collaborative group level type of production process. Two generations ago, almost all of the economics papers in top journals would have been sole-authored, one person working by-- well, mostly himself, occasionally herself, to produce economic research. A generation ago, it would have been half sole-authored and maybe have two authors. And now there are these big interdisciplinary research teams. And it's really just an absolutely extraordinary change in my time in the profession.

SARAH HANSEN: I mean, that's such helpful context. And I love the vision and the practice of where it's going, because it really shows that so many people can be involved with economics. And so many people can contribute from their various positionalities, and that there is room for many, many people and perspectives, which I think sounds like it's really important for moving the field forward.

So you two have collaborated as well as educators. And that's its own unique form of teamwork. You came together to teach 14.310x, which is Data Analysis for Social Scientists. Esther, could you just share an overview of what the course is about?

ESTHER DUFLO: So the ambition of the course is really to cover everything from probability to machine learning. So this is often done separately in different courses. Often, the probability part is cut short and you start with statistics without any notions of where the statistic is coming from, which does lead, in our opinion, to some half-baked thinking about what statistics is really representing, or what statistics is used for and how. And we wanted to give people with not necessarily a huge background, but definitely a willingness to hold on to the train to equip them in the course of a semester to be ready to carry out their own work, in particular on empirical work, across the social sciences.

SARAH HANSEN: And then Sara, could you give us a few examples of the kinds of social programs that students are interested in evaluating?

SARA ELLISON: Sure. I mean, it really runs the gamut. Let me take a step back and just say that there's a set of-- if I can call them, legacy tools from probability, statistics, and econometrics that are part of the standard tool kit of empirical social scientists. So like for instance, if we go back to the example that Esther mentioned a couple of minutes ago. You're interested in the effect that introducing computers into schools might have. She described in some terms about how you would maybe randomize 200 schools. 100 of them get the computers and 100 of them don't.

And then you would get data back in lots of different forms. You would get data on how the students are performing in these two sets of schools. You would get data on how the faculty liked the computers or didn't like the computers. You would get data in lots of different topics and lots of forms.

And you have to know how to rigorously analyze these data. So you need to be able to look at them and say this is a difference that could have just happened by chance statistically. Or this is a difference that with very, very small probability could have happened by chance if there was no actual difference between the two sets of schools. And so those are the kinds of legacy tools that we teach students learn.

And then the course goes beyond that. And we talk about issues like causality, and we introduce some machine learning techniques, et cetera. But the legacy tools that we focus on, those are just a standard part of any social empirical social scientist's toolkit and what they would need to use to analyze any policy.

SARAH HANSEN: OK, interesting. Could you speak specifically to the role of AI that is emerging? It's at the front of everybody's minds. What role is AI going to have in the evaluation of social programs when thinking about economics?

SARA ELLISON: Yeah, my sense is that there are just lots of examples cropping up in the last few years of resourceful social scientists using AI and ML tools-- machine learning tools to enhance or aid their research in various ways. So one example might be relying on AI to do things somewhat automatically that you used to have research assistants do manually. So just to take one example from my research,

I'm working on a project on the effect of digital platforms on the real estate market in New Hampshire specifically. We have a data set on every real estate transaction that's happened in the last 20 years in the state of New Hampshire. And we want to classify these properties that were transacted as being waterfront or not waterfront.

And you can imagine, if you were doing it manually, that would be impossible. And there are AI tools that allow us to automate that process in certain ways. So that's just one small example. But maybe Esther has others.

SARAH HANSEN: Yeah, Esther, I'd love to get your take on the role of AI in data analysis and how you see that shaping the field, and maybe any ethical concerns you might have.

ESTHER DUFLO: So I don't know about AI per se for evaluation, but machine learning for sure. AI seems to be a bit of a grand word for a lot of what machine learning is doing, which is basically estimating-- with data sets, it's large. It just happens to be data set with a lot of variables.

So you might think that randomized controlled trials and machine learning are not really compatible because most of machine learning's objective is to do prediction. Given a set of things you know about, for example, an individual, can you predict what they are going to do next? So the large language models are very much that. Given all of the words that have been being said, after these first three words, can you predict what comes next? So AI or machine learning models are fundamentally at heart prediction exercises.

And randomized control trials are almost the opposite, because they are meant to be causal inference and very single parameter-- usually one, which is what is the effect of some variable y on some variable x. So you would think there is very little in common. But in fact, that's not true. So a lot of my work on the more theoretical side lately, in collaboration with machine learning specialists in our department-- in particular Victor Chernozhukov, has been to combine the two to make experiments more powerful.

So for example, you might ask yourself, what type of people are going to be most affected by your intervention. And then you're back to a prediction model because you're trying to predict the type of people that are going to be the most affected by an intervention. It's a prediction about a treatment effect, but it's still a prediction.

It doesn't necessarily have a causal interpretation, but it can still be useful to guide policy if it's done well, because it could tell the policy maker, well, you have this intervention, say, to improve immunization rates. Maybe it's going to work better in these type of regions than in those type of regions, so you can focus the program in the regions where it works the most. So that's an example where you can use the two together.

Another place where AI is starting to play a role, or at least has the potential to play a role, is in intervention itself. So for example, I'm working with Ziad Obermeyer from Berkeley, who is both an MD, and a statistician, on developing a prediction model to find out where older people in the south of India have had a heart attack in the past. So when someone has had one heart attack, they become more susceptible for the next one.

But there are many things you can do to really lower that risk if you know that it has happened. But in a lot of cases, you don't know, because if they didn't die from the first heart attack or they weren't hospitalized, then maybe it went, they felt ill for a bit of time, and then they go back to run their lives. But it leaves scars, which are detectable by very expensive tests for a lot of rural people in India. So our objective is to see whether we can develop a prediction based on tests that would be much cheaper to do. So what we have done is invited 6,000 people to health camps where we tested them with the best tests that there is-- ECG, ultrasound, retina scans, and so on, as well as things that are much cheaper to do.

And we have just finished collecting the data. And our hope is to then train a model that helps classify people into green, yellow, and red. With red, this person needs to be referred because they are at risk.

Or this person needs to be monitored in orange. Or this person never had a heart attack, had a great heart, and we can leave it at that. So that would be one example of using AI for good in these settings.

SARAH HANSEN: Yeah, thank you for that it's a wonderful example. And I think something you said earlier is, it's actually really important for laypeople to understand that ML and AI is largely predictive. It doesn't get at causation. And I think for people who are not in the field, it's easy to take something that AI says as true, unless you're really thinking deeply about what these distinctions mean. So thank you.

I'd like to talk about the unique aspect of this course, in that it's offered on campus, but I believe it's also offered as a MOOC, or a massive open online course. And Sara, I was wondering if you could speak to that a little bit. Like, what is the MOOC offering? And how is it connected to the MicroMasters program?

SARA ELLISON: Yeah, happy to talk about that. So in fact, we did turn this course into a MOOC. And it is available online for anyone who has an internet connection.

They can take the entire thing for free. We do charge if you want to take the proctored exam to get a certificate saying that you took the course and passed it. But all of the material in the course-- and this includes problem sets and exams and access to teaching assistants and online teaching assistants, et cetera, is all available free to anyone.

So you referred to the MicroMasters. We have a MicroMasters program, which is a credential that learners can earn if they pass five courses in a suite of courses that we have that's related. And this is one of the courses in that suite of courses.

And the focus of the MicroMasters is on data, economics, and development policy. The data analysis course that Esther and I taught is just a really important piece of that, because in order to think about development economics, up to policies that might address questions of development and be able to evaluate those policies, having the data analysis toolkit is a prerequisite. It's an indispensable part of that.

SARAH HANSEN: Wonderful. And Esther, I'm curious if you could tell us a little bit about what made you want to pick Sara to be your partner in this. What do you value about her as an educator or as a researcher?

ESTHER DUFLO: It would take a very, very long time to speak about everything I value about Sara. But simply, put she's the best professor I've ever met. So it's hard to think of a better person to teach this course or any course. And in particular, it's material that some people would find hard to grasp or would be a little bit scared getting into.

But the delivery by Sara is so clear and so devoid of magic that you immediately get in. And when we were teaching, in the class, we used to go to each other's lectures, and I feel I understood more probability than I had ever before. So I'm hoping that's true with students who have had a little less experience than [INAUDIBLE].

SARAH HANSEN: Wow. Well, so for anyone listening out there, sign up for this book because-- what a wonderful opportunity. I'm just curious about this thing you say in the course description, which is that students learn how to present results in compelling and truthful ways. And Esther, I'm just curious, is it hard to be both compelling and truthful?

ESTHER DUFLO: No, it's perfectly possible. But it requires a little bit of work.

SARAH HANSEN: Tell me more.

ESTHER DUFLO: In the sense that it is very easy to present data visually in a way that distorts the truth. And there is actually a whole field of data visualization that tells you what to avoid to avoid doing that. It's not always conscious. Sometimes it is.

But often, it's not conscious. For example, when you're trying to do a bar chart, Excel will often default to not start the bar at 0. So if you have, for example, a treatment effect that moves a percentage of people who decide to get an immunization from 90% to 95%, it is a difference, but it's not a humongous difference proportionally. But you might make it look really, really big if you start your axis at 92%, for example, because then your first bar is going to be very low. And the second bar is going to be very high.

And our eyes, as a reader, will think of the proportional change to be from the small bar to the big bar, even though it is written that it starts at 92%. So it is misleading to start a bar chart from somewhere in the middle. But it is offered by default in Excel.

So I don't think it is out of my list. So there is a number of things like that that one has to think about. Always start your bar chart at 0 is one of them. And through these days, I have to tell my students, your bar chart is beautiful, but if you start at 0, it's not going to be as impressive.

SARAH HANSEN: Yeah. Well, I mean, this is so important for consumers. I'm not sure that everyday people are thinking about that when they're looking at data visualizations. So I think it's really important for students to learn about this, so that when they communicate their science, it's as truthful as possible. But it's also important for consumers to be looking for things like this and to question data.

ESTHER DUFLO: It is very important, because often, of course, there are people who are in the business of impressing, and not necessarily in the business of truthfully communicating. And therefore, they will actually look for ways to present data in a way that's not wrong, but that's misleading. So it's also useful for readers. In any statistics class that should be part of the course, to get some of these pointers about how visualization can be deceiving or truthful on the other hand.

SARAH HANSEN: Especially when we're looking at social programs and the impact on real people's lives, this is just critical, I think, which is also why this course, is really important. I'd love to hear from both of you if there's anything you would like to change or add to the course in future iterations. Sara, I see a thought bubble above your head. Would you like to go first?

[LAUGHS]

SARA ELLISON: Yeah, I guess I would say that I would love to build out additional scaffolding that goes along with this course. So as I suggested before, there's a lot of material in this course. It's dense, it's difficult, it's challenging.

I mean, we think that's good. But there are students who approach the material, and they could really benefit from some extra scaffolding. And part of what I did in the past couple of years is wrote a companion textbook for the course. And so that provides some of the scaffolding, but I would like to work on even more support materials as well.

ESTHER DUFLO: And if I can add to that, I was hoping that Sara would mention that, because I think that that's exactly one of the things that has already happened. And one of the things that is also happening is, too, we are trying to work with institutions around the world for them to use this material as part of their teaching for their own students to provide this scaffolding in classrooms to sessions that can be done in complement to the classes. Often, some people might need that either for moral support or for technical support or both. And over time, what we are doing, and what needs to be done even more, is to develop tool kits for the faculty who are teaching these kinds of supporting sessions.

SARAH HANSEN: That's really exciting. There's nothing more powerful than open resources being used in context and supported by educators to meet the needs of specific learners. That's when it all takes off. That's the ignition point.

ESTHER DUFLO: So the two of us are about to go to Cote d'Ivoire, where the statistical school there-- Statistical Institute is adopting this class and other classes in the MicroMasters, but particularly this class and the basic micro class as part of their teaching.

SARAH HANSEN: That's fantastic.

SARA ELLISON: And I just looked up-- by the way, the weather in Abidjan is 90 degrees these days.

[LAUGHTER]

[INTERPOSING VOICES]

ESTHER DUFLO: That's not why we go.

SARA ELLISON: That's not why we go.

SARAH HANSEN: Well, congratulations, to both of you. And thank you so much for being here. It was delightful to speak with you.

I learned a lot. I think our listeners will learn a lot. And I hope a lot more people take your MOOC because it sounds absolutely critical for consumers, and people just even curious about how to get involved with economics.

ESTHER DUFLO: Thank you so much for having us.

SARAH HANSEN: That was Esther Duflo and Sara Ellison, co-creators of course 14.310x, Data Analysis for Social Scientists, which you can find on MITx and OpenCourseWare. As always, their materials on MIT OpenCourseWare are openly licensed so you can reuse and remix them in your own teaching and learning. You can help others find the materials too by subscribing to the podcast and leaving us a rating and review.

Thank you so much for listening. Until next time, signing off from Cambridge, Massachusetts, I'm your host Sarah Hansen from MIT OpenCourseWare. MIT *Chalk Radio's* producers include myself, Brett Paci, and Dave Lishansky.

The show notes for this episode were written by Peter Chipman. Cheryl Siegel built the course, on our website. Jason player made our episode cassette animation on YouTube. We're funded by MIT Open Learning and supporters like you.

[THEME MUSIC]