

[SQUEAKING]

[RUSTLING]

[CLICKING]

**FRANK**

Welcome to Lecture 18 of 14.13. This lecture's about gender discrimination and identity.

**SCHILBACH:**

OK. So what's the agenda today? Generally, we're going to talk about gender, identity, and particularly, discrimination. First, I'm going to give you a broad and a quick overview of why this is an important area to study. Then I'm going to give you an equally brief overview of the gender gap in wages, the fact that women earn less money than men for equal work, how that has evolved over time, and how much still is left of this gender gap.

Then we're going to talk for a little bit about potential technological solutions. Can we use technology to improve or to reduce discrimination potentially? And the answer would be yes, to some extent. But that technology will not solve everything.

Then we're going to talk about two particularly important issues. One is beliefs and people's updating about beliefs when it comes to men and women and their skill. And then we're going to talk about gender identity norms in particular, which is the idea that people's deeply held beliefs about what should be done in society or not might affect their behavior.

And finally, we're going to talk a little bit about demand and supply for different tasks. So in particular, I'm going to show you that women are more likely to say yes to what's called nonpromotable tasks, tasks that are good for society, but perhaps not good for their career, or tasks that are good for, like, a company or a university as a whole, but are not good, are not promotable in the sense that they help women, in fact, get promoted. Women are more likely to say yes to those things. In addition, on top of that, precisely because they are more likely to say yes, they're also more likely to be asked in the first place, which sort of amplifies that difference. I'm going to tell you about this in more detail.

The first question you might ask is why study gender differences? I have four answers for you. The first one is simple equality, fairness, and justice issues. People simply should be rewarded equally for the same output.

So if a man and a woman works equally many hours and is equally productive, they should be paid the same. And if that's not the case, that's unequal. That's unfair and not-- unjust, or not justified. Now there might be additional equality and fairness issues that men and women might not even be allowed to work as much, women might not be allowed to work as much as men. That's an additional sort of fairness concern.

Second, there's an efficiency argument, which is overall productivity and welfare falls if women and other groups are held back by discrimination and other distortions. And indeed, a substantial share of recent US growth can be explained by improved allocation of talent in the economy. So this is just a simple idea that, well, if we exclude certain groups, including women from certain jobs, the most talented people in those groups will not be allowed to-- or some of the most talented people in those groups might not be allowed to do those jobs. And that's bad for society overall.

Think about, for example, doctors. If women cannot be doctors, well then, we will essentially just exclude a large fraction of the most talented people in terms of being doctors, the smartest, brightest and hardest working women. And that will be worse not just for those women, but it will also be bad for medical care, innovation, and so on in the medical profession because you just essentially exclude some of the best people from that type of work. And if you want to learn more about this, the paper by Hsieh et al. is an excellent discussion of this issue.

Third, we might learn about the formation of preferences in personality by studying gender. And here, the idea is that there are, in many cases, differences and preferences-- [INAUDIBLE] preferences, social preferences-- and taste for competition or competitiveness but also attitudes towards negotiation, or other things such as identity aspirations, or over- or underconfidence more generally.

Well, so if there are these differences at later ages, we can learn about the formation of those preferences in personality by looking, for instance, at people at different ages and trying to understand at which ages those differences emerge. For instance, you might look at 5-year-olds or 10-year-olds or 15-year-olds and look at their risks and social preferences, and then try to understand when those gender gaps emerge over time. And that will allow us, then, to understand questions of nature versus nurture, whether there are inherent differences in preferences and personality overall, just men and women are just different inherently for some reason.

Or it could be that it's really about socialization and education and the like. Girls and boys are just treated very differently by their parents, by their social environments, their teachers, and so on, and that might create such differences and preferences and other behaviors. So that's another reason to study those. You're going to mostly focus on number one, just simply the idea, the objective that people should be rewarded equally for the same help.

Now what is the gender gap? What do I mean when I talk about the gender gap? So Claudia Goldin, who is a eminent researcher at Harvard, who has done a lot of seminal work on gender and the economics of gender, has written a very nice overview paper in 2014 and giving an overview of what do we know and how far have we come in trying to reduce and close the gender gap. The graph that I'm showing here summarizes this discussion fairly well.

It's a bit of a messy graph. So let me show you what it shows, or tell you what it shows. On the x-axis, you see people's age from 20 to 70. On the y-axis, you see the log of female to male earnings, or average female to male earnings. Which is essentially a measure of the gender gap, and so the difference in log is essentially the percentage difference in female versus male earnings. So minus 0.3, and we think of that as, like, 30% difference in female versus male earnings.

What do we see here? Well, A, we see a pretty large gender gap anywhere in this graph. I should have also said different lines are now different cohorts. So these are starting in 1923, so people born in 1923, '28, '33, '38, '43, and so on, at five year intervals, from the oldest to like more recent cohorts.

Now what do we see? We see a significant gender gap everywhere in this graph, or in all lines of the graph. We see that the gender gap, even for the most recent cohort seems to be about 20% to 30%. We see that the gender gap has fallen. So when you look at like the most-- the oldest cohorts, the gender gap was even larger than the more recent cohorts.

You also see that there's sort of this U-shape, where we seem to find that the gender gap starts emerging already at age 27, 28, 30. There's already a gender gap of something like 10% to 15%, even for the most recent cohorts. But then that gender gap gets magnified towards, like, [INAUDIBLE] 20% 30%, 40%, until about age 40 to 45, when it's at a plateau. So maybe it falls a bit overall.

So what do we learn here? Well, there are substantial female labor market gains in the last half century. But large gender gaps remain.

In 2010, for instance, the ratio of mean annual earnings between male and female workers, these are people who worked full-time, full year, 25 to 69 years, was 0.72, and that of the median was 0.77. Notice that the figure here controls for work time and education. So while of course, it might also be the case, it is indeed the case that women work fewer hours in a day than men do, even conditional on work time education, there's a significant gap, gender gap remaining of something like 20 to 30% that women are paid less for equal work.

Now first, what are some of the reasons why there have been substantial female labor market gains over the last half century? Well, there's quite a few technological and other explanatory factors. One big factor was a reduction in the gender gap in education. Women are just more able and can afford to receive education. That's true for both primary, secondary education, but also, in particular, for university or higher education, more generally.

Second, there's been technological innovations, such as the pill, dishwasher, and so on, which allowed women to do more work outside of the household. There have also been labor demand shifts. In particular, the US economy has been moving from manufacturing and agriculture towards services. Women tend to have comparative advantages in services compared to like, in particular, hard manual. And so these labor demand shifts then have disproportionately benefited women.

Finally, there's been also lower discrimination, in particular, stronger regulatory controls and increased market competitiveness. For the latter, it's if there's more competition, firms cannot afford to discriminate as much. Because it's actually a costly thing to do, to hire incompetent men compared to competent women.

Now we're going to discuss a paper by Goldin and Rouse, which is an example of a simple technological solution, in particular, this question whether blind auditions in orchestras increased gender ratios in those orchestras. So is the fact that people-- the blind auditions were introduced, did that contribute to the increase in the gender ratio in orchestras?

What do I mean by the increase in gender ratios in the orchestras? So even until the 1960s, '70s, and '80s, there were shockingly low female ratios in top orchestras and orchestras more generally. Now these ratios were below 10% in many orchestras. And only recently, in the '80s and '90s, in particular, there were stark increases in the number of women who play in those orchestras.

This increase was the case in the top five orchestras, but [INAUDIBLE] also has been, have been similar improvements in gender ratios in other orchestras, which you can read in the report that I linked in the slides.

Now one really important reason here seems to have been blatant sexism, even by very renowned, or in particular, by renowned conductors, which you can-- I'm not going to repeat here, but you can read it in section one of the paper by Goldin and Rouse.

And so one question is like, well, is there's this sexism and discrimination, has that contributed to the low gender ratio in orchestras? Now while there have been all these improvements in the gender ratio [INAUDIBLE] musicians in general? There still are vast differences in conductors and the number of conductors and music directors. They're essentially predominantly male still. Moreover, while there have been a lot of improvements in terms of gender, most musicians are still white. So in particular, African-Americans and Latinos are very much underrepresented in orchestras.

Now let me tell you a short aside about women in economics. This is not just the case that in orchestras the fraction of women is low, that's also true in economics. And here's an overview of that. There's a [INAUDIBLE] paper by Lundberg and Stearns in 2019, which essentially says that some progress has been made, but lots more needs to be done. And more recently, in particular, some of the progress seems to have stalled.

What do we mean by that? Well, here this graph shows the fraction of women in different stages of the profession ranging from senior managers here at the top to first-year PhD students to assistant professors, new assistant professors, associate professors, and full professors. And what you see here, there seems to be some trend that's upwards, in particular for full professors and associate professors, which is, of course, good. there's less of a positive trend to be seen in particular when it comes to assistant professors.

Here you essentially see since 2000-- about 2007 or 2006 has been no increase, if anything, like a decrease in the fraction of female assistant professors. But even the fraction of PhD students is clearly below 50%. And so the profession is starting to understand and address the issues, in particular, to try to reduce sexism, and other issues holding women back.

For example, some particularly nice examples of these mentoring programs, which are trying to support female and minority assistant professors and in terms of mentoring and trying to sort of foster their careers, which has been shown to impact and significantly improve the profession. For what it's worth in '14, '13, the majority of students tends to be female. And moreover, the best students tend to be female as well. So very much hope you are still interested and all the female students are very much still interested in economics and would pursue a career in economics, perhaps academic.

Now back to orchestra, so many orchestras introduced blind auditions in the 1970s and 1980s. And if you look back at the graph that I showed you-- this graph here-- you see in the 1970s, '80s, and '90s, that's also the time when there was a sharp increase in the share of women in those orchestras. So you can see here in the graphs, orchestras, if you want to get a job in orchestras, you have to do auditions.

These auditions are very, very competitive and often have like several stages. There are preliminary, semifinals, and final symphony orchestras. What you see here is 11 anonymized orchestras, including the top orchestras in the US, and the introduction of blind auditions in those orchestras.

What you see is that most orchestras did introduce blind auditions at least in some of their stages of the auditions. However, many of them also only introduce it in some stages, but not in others. Now you can think already that that's not sufficient.

In some cases, if you have only like blind auditions in the preliminaries, but not in the semifinals and finals, you might sort of reduce discretization in the preliminaries. But then if women are discriminated in the semifinals or finals, that's not going to help them get hired, which is, at the end of the day, we, of course, care about.

Now Goldin and Rouse look at technological solutions and look at the question of whether such technological solutions, in fact, can help. Now the question to ask now is looking at data from actual auditions, often using individual fixed effects, do these blind auditions, the introduction of blind auditions help candidates reach the next round or, in particular, get hired.

Now the data is fairly rich and often includes or allows for the inclusion of individual fixed effects. That is the case if you look at the musicians, Goldin and Rouse are going to ask the question, is a particular musician more likely to make it to the next round when auditions are blinded? And is that more likely the case for female than for male musicians?

Now what did they find overall? They find evidence that blind audition procedures fostered impartiality in hiring and increased the proportion of women in symphony orchestras. Now there's some caveats here that, you know, some of these estimates are fairly imprecise, in particular, in the later stages because a number of observations was down quite a bit. Some of these results are a little bit mixed. But overall, the evidence suggests that these blind auditions did indeed foster impartiality and help women succeed in getting jobs, in particular in these top orchestras in the West.

Now why are these results perhaps somewhat less clear cut? I already told you like one reason is that in the later rounds, the number of observations gets kind of small because no longer have so many people make it to the later rounds. But there's another reason, which is the auditions have several rounds, not all of which were blinded. Now if there's one-- if there's discrimination in one round, but not in the next one, this issue that in the next round, there will be like selection, and in particular, the discriminated people who make it to the next round, you expect them to do better, right?

If there's discrimination against women in the first round, but not in a second, you think that women in the second round will do better because only the best women will make it to the second round because otherwise they will be screened out in the first round due to discrimination. But if that's the case, then making comparisons in later rounds is potentially biased or problematic because you're comparing essentially different types of people in different rounds. And that leads to the results being less clear-cut overall.

Let me give you another example. Suppose male and female workers get graded on their task, a coding task, online where you essentially see their output, but you also see their gender. Now suppose women get graded systematically worse due to sexism and no other reasons. Suppose women and men are exactly the same in the work they do, but some men are sexist and they're going to rate women worse. Now for any given score, you should expect the women that you hire compared to any other random man that you hire to do better than that man. And the reason being that they have been systematically disadvantaged due to the sexism in the rating to start with.

Now you should think about this. Because, you know, there's also some evidence of teaching evaluations being biased against women. Now if you see a man and a woman are at the exact same teaching rating, well, you would probably want to take the woman's class because the woman has been-- the female professor has been discriminated to start with. Having that high score or the similarly high score as the male professor is more impressive and likely to indicate that the class will be better.

There's an excellent paper by Bohren et. al. that considers the issue of-- the dynamics of discrimination in much more detail if you're interested in learning more about that. Now despite the technological solution that advances that image, [INAUDIBLE] substantial gender gaps remain. So while, as I told you, the gender gap has been reduced both in terms of labor market or labor force participation, how many days and hours, how many women overall, how many days and hours those women work and in their wages and earnings, how much they are being paid conditional on working, there are still substantial gender gaps.

In particular, women's labor market participation has plateaued since the early 1990s. And even now, among entering cohorts, women still earn significantly less than men. And that's true even for conditional on work time and education. That's the graph that I was showing you earlier.

Now given that there are these persistent remaining gender gaps even though a lot of the technological and other barriers have been reduced-- now women are educated or as educated as men are and there's less technological barriers that preclude women from working, researchers have now started to consider some of less traditional within economic factors, including things such as risk attitudes, negotiation skills, taste for computation, but also-- and this is what we're going to focus on-- beliefs and social norms and identity. I'm going to first talk about beliefs and then we're going to talk about social norms and identity.

Before we get to beliefs about gender, people's-- women's and men's skills, I'm going to tell you very briefly about a seminal paper by Bertrand and Mullainathan from 2004 that randomizes-- randomized names in job applications. So what they're trying to do-- they essentially sent out job applications that were otherwise identical, male and female workers. These are fictitious job applications. And they sent them to employers and then looked at callback rates by those employers. And the only thing that was like varied in those applications was the name on the application.

And what did they do? Well, they used white-sounding names such as Emily and Greg, as well as African-American sounding names such as Lakisha and Jamal. And then they looked at, well, were these applications, when they had white-sounding names that were otherwise identical, more successful than applications by names with African-American-sounding names and, again, that were otherwise identical?

And Bertrand and Mullainathan find a striking result that callback rates for white-sounding names were 50% higher than for African-American-sounding names. That's a huge difference. And that's important not just because it's a hassle because you have to send out more applications, but if you get called back more, you're more likely to be interviewed or asked for an interview, you're more likely to actually get the job, you're more likely to-- less likely to be unemployed, and you might also be less likely to stop searching whatsoever because you might be disappointed or discouraged.

So that's a huge difference and, essentially, discrimination against African-Americans was blatantly obvious from this research. Numerous other studies show gender and racial biases in various ways. These are ranging from very similar studies, including these audit studies where people are sent resumes and the like, but also ranging to studies where like people on eBay tried to sell things, where they try to sell like an iPad or iPhone, and they hold that iPhone with hands that's either like white or non-white.

And then they find essentially that white hands are more likely to sell their iPhones than non-white hands. Now this is a very important and depressing result. One thing to flag here that it's hard to distinguish statistical from case-based discrimination. You're going to talk more about this recitation.

What do I mean by that? Well, it might be that there is-- people have a preference for white applicants. That is to say, controlling for performance, it might just be that people are racist and they'd rather have white applicants than African-American applicants. This is what economists would call case-based discrimination, which essentially is conditional on the performance. People just like certain types of people better than others. And, therefore, they call them or are more likely to call them back.

A different explanation is belief-based explanation. It might be that employers think that African-American applicants will perform worse, even controlling for all other aspects of the resume. So it could be essentially you see, people see this name and they think that even if the resume is good, that person will not perform work as well as somebody with a white-sounding name. Notice that these beliefs could be correct or not, but in any case, we will-- and that's-- we'll talk about this in recitation as well. Economists again, would call this statistical discrimination.

This could be, again, correct statistical estimation or statistical discrimination that's based on biased beliefs. But either way, these things jointly led to clear discrimination of African-American applications, which, in turn, leads to striking differences in unemployment or job finding rates across races.

We're going to now focus on beliefs and, in particular, about beliefs about skills and how people update on those beliefs depending on the gender of the person they're updating on. The people we are going to consider now is the paper by Sarsons, 2019, which considers how people interpret signals in the labor market. The question that Sarsons asks here is whether somebody's gender influences the way we interpret information about this person and about his or her peers. What do I mean by that?

Well, in many situations when people do certain tasks and perform at work, good things and bad things happen. Now you can infer from these things that occur, from these events, you can infer something about their competence and their performance. While many of these situations are ambiguous in the sense it could be that the person was just really lucky or it could be that the person is really, really good at what they're doing. Now what Sarsons is asking when we see those events and when people see those events, are people interpreting those events differently for men versus women?

So, for example, if a man-- if something bad happens at work, a boss might react to that person differently or update their beliefs about the person's skills differently depending whether it's man or woman. In addition, the boss might also update differently about the peers. So if a woman makes a mistake, the boss might not only update his beliefs about this woman, but also about all other women that work for him.

Similarly, for a man, of course, the boss might do the same. And what Heather Sarsons now asks is the question, is this updating or the interpretation of information differently for male versus female workers? And why do we care? Well, we care a lot about this because hiring, promotion, and wage decisions hinge on information about workers' ability. And in particular, if there's systematic difference in how information about men and women is interpreted, that could lead to differences in hiring, promotion, and wage decisions and, in turn, might contribute to the gender gap.

What does Sarsons do to answer this question? She looks at how do physicians change referrals to surgeons and their peers after a patient, after certain patient outcomes? Now what I mean by that, physicians often refer patients needing surgery to a local surgeon. So somebody breaks their arm or their leg, well, many physicians can't do this on their own, so they will refer this patient to a local surgeon.

Now the referral choice reflects the physician's beliefs about the surgeon's ability, right? You want to do the best for your patient. And you want to send that patient to the best possible surgeon. So if you have a suspicion, if you think a certain surgeon is not competent, you better not send your patient to that surgeon, in part, because you might care a lot about the patient, in part, because, you know, if you make bad referrals, that might also reflect on you eventually.

Now what does Sarsons do specifically? To document whether the reaction depends on the surgeon's gender. What does Sarsons do? She matches on surgeon and patient characteristics on procedures. That is to say, she takes male and female surgeons and patient characteristics and procedures-- which is like, you know, broken arms versus more complicated procedures-- so she matches all of those characteristics and such that in observations set, she has surgeons only differ by gender. And then she considers what we call an event study, which is comparing how physicians reacts to male and female surgeons when good and bad things happen.

Let me now show you more precisely what we mean. So we have here on the x-axis, quarters. On the y-axis, we have the referrals from the physicians to the performing surgeon. So these are essentially dyads or pairs of physicians and performing surgeons. And you see here minus 1 is normalized to be 0. And then, essentially, you see the evolution of the referrals from a specific physician to a performing surgery.

These are men and women. And these are essentially cases in which there is no adverse event or the adverse events are like a patient's death, for example. What you see are essentially these are men and women that are matched. They're supposed to look exactly the same. And what we see here is that referrals increase over time. They tend to sort of increase fairly steeply over time and then sort of plateau off over time.

The exact shape and pattern of this relationship doesn't matter so much. However, what matters is that men and women look exactly the same, right? So the performing surgeon is here. The gender here that varies is the gender of the performing surgery.

Now then she looks at-- Sarsons looks at cases in which there is an adverse event. Again, the adverse event happens around here. So what you see here is then when there is an adverse event, there's a clear punishment in the sense of like the male surgeon now gets fewer referrals from the physician right after the adverse event happens.

So this is manifested by the blue bars being lower than the gray bars. Notice that the number doesn't go down. It's still above-- everything here is above the red line. So it's just the physician is not increasing the referrals over time anymore or that increase gets dampened by the adverse event. This is for the male surgeon.

Let's look at now what happens to the female surgeon. What we see here is a clearer difference that opens up after the adverse event. Notice that there's no pre-trends here. Before, it looks-- men and women, or female and male surgeons, including the red and the blue bars look exactly the same. But the bar, the difference only opens up after the adverse event.



And there's a pretty large gap, you can see, that shows up in the referrals. And this gap is way larger if you compare the red versus the gray lines-- it's way larger compared to the gap with the divergence between the blue and the gray lines. So what we see here is that both men and women are punished in the sense of receiving fewer referrals after an adverse event. That punishment is way more severe for female than for male surgeons.

Now let's do the same for the absence of a good event aren't actually good events that happen. What's a good event? Well, it's essentially a surprisingly good thing that happened where there's a complicated surgery and the person does really well. What we see here is, again, sort of the men and women, they look pretty much exactly the same. You can see here maybe men are doing slightly better overall.

But overall, these gray lines look very similar over time. Now, if there's a good event that happens, you see, essentially, men get very much rewarded. If the male surgeon does well, he gets rewarded in the sense of receiving a lot more referrals from the PCP, from the doctor. Now if you do the same for female surgeon. Again, female surgeons benefit from that.

But the increase in referrals is a lot lower than for men. So, again, both men and women, there is a reaction to that information, to a good or a bad event. But the reaction tends to be much stronger, much more favorable for men. In terms of the good events, men get rewarded more for the good events. And women get punished more for bad events, [INAUDIBLE] bad events that happen.

Let me summarize the main results. After a bad outcome, after, for example like a patient death, there is a 34 decrease in referrals to female surgeon-- surgeons. And there's a stagnation in referrals to male surgeons. So, essentially, women tend to be punished quite a lot. Men, a little bit in the sense that referrals don't go up anymore, but the punishment is much less.

Now, strikingly, physicians are also less likely to refer to other female surgeons. That is to say, if a female surgeon, if a patient dies from a female surgeon, the physician is also less likely to refer patients, future patients to other female surgeons. These other female surgeons, nobody has died there. So, essentially, what seems to be the case is that the physicians also are updating about women or female surgeons in general if there is a negative event that happens, which is sort of very striking.

After a good outcome, such as unanticipated survival, there's a doubling of referrals to male surgeons but only a 70% increase in referrals to female surgeons. And there are here-- there's no spillovers to other female surgeons. So, again, what do we find here? Well, we find that women are punished more for unexpected bad events. And they're rewarded less for unexpected good events.

That is to say, it's not just the case that there's more updating going on, in the sense maybe you learn more from good and bad events because maybe you have like less information about men and women to start with. That's not the case here. What is here the case is that there's more updating or more punishment for mistakes for women and less reward for good things to happen, for good events for women as well.

Now, why do these asymmetries matter? Well, one important issue here is that women have more chances to make mistakes. And, in particular, if there is like dropping out, if women don't get promoted or if they even drop out, that leads to lower skill accumulation and, you know, women will be just less qualified overall in the end because they have less chances to get promoted and get more responsibility and so on and so forth and then they even drop out overall and we just have fewer women in the profession.

Now, moreover, usually wage gaps are measured conditional skills, industry, and position. But if there is biased evaluations that then lead to differences in skills and positions in the workplace, not only is it the case that women are receiving wage gaps conditional on the skills, but they also have different skills and positions in the workplace because of biased evaluations. So that's an additional scope for discrimination and differences in earnings and so on.

And then, importantly, this is also not a mistake that can get corrected over time. Because if women drop out in particular, then-- or if there's no referrals that happen anymore, then the doctor or any other person might not receive any signals anymore from work. If you sort of update on a particular person that that patient of mine died because I sent them to that person, if you then think, you know, this person is terrible and don't send any further patients anymore, you will never learn that this woman was, in fact, very much highly qualified and that was just an accident that could have happened to anyone.

So that's a really important issue that-- so when women are underrepresented employers see fewer outcomes overall, and if there is updates more of the bad outcomes, or, in particular, if there's also updates about all women, then there will be fewer chances to learn and then women are let go and so on or less likely to be hired, then there will be no chance of actually correcting those [INAUDIBLE].

Strikingly, as I told you, what's really important as a result is that seems to be like updating, not just about one specific woman overall after the negative events, but also about all other women that get punished for the bad event that happens. And if that's the case, if there's several women in the profession, then all women always have to pay for any bad events that happen to any woman. And sometimes, you know, mistakes, in fact, happen. That's, of course, terrible because that sort of amplifies a lot of the updating. Because even if somebody makes no mistakes whatsoever, negative updates happen on their performance, other women perform negatively in some ways.

Now one question you might ask is, well, couldn't we just use algorithms, machine learning, et cetera, to overcome such biases? In particular, instead of hiring and so on, couldn't you use algorithms and machine learning to train those algorithms and instead of use computers instead of humans and then, you know, we might get rid of those biases? Well, one key problem with that is that algorithms themselves can be biased. Now why is that? Well, we train algorithms based on human decisions.

As an example, if you train an algorithm-- this is, in fact, what some companies did, based on actual decisions, you get asked the question, this [INAUDIBLE], if you train that as an actual hiring decisions or interview decisions or you say, well, would this person be invited or is this person, is this resume likely to be invited for an interview? Well, if those interviewing decisions to start with were biased based on sexism, racism, and so on, if you train your algorithm on it, then very much you make that algorithm biased as well.

Now, crucially, that's, of course, problematic and I very much would like to avoid that. And we want to be very careful. But, crucially, some of these biases can be perhaps easier or fixed and perhaps fixed more easily than biases in human decision-making, right? If somebody is sexist or racist, it's very hard to fix this, at least in the short run. However, an algorithm, essentially, is just doing whatever you train them to do. So if you build equity concerns into the objective function of an algorithm, then we can essentially reduce or even eliminate these biases explicitly because the algorithm is explicitly asked to do so.

And so there's a fascinating-- if you want to learn more about this overall, there's a fascinating talk on discrimination by algorithms and people by Sendhil Mullainathan that I've linked here in the slides that you can watch it and learn more about. But overall, so just to say, to summarize, algorithms can be biased, but they can also be fixed. And that's an important research agenda that people have been working on more recently, the design of fair algorithms.

Next, we're going to talk about gender identity norms and a very nice paper by Bertrand et. al. from 2015. Identity considerations were imported from social psychology to economics in multiple papers by Akerlof and Kranton in particular in 2000. There's a very nice book about identity economics if you're interested in learning more about that. Gender identity norms are an important example or application of such norms or identity considerations more generally.

In particular, Akerlof and Kranton, or we here as well, focus on two social categories, men and women. For simplicity, we can also just talk about husbands and wives, understanding things are more complicated in reality. But for simplicity, we're going to focus on that for now.

What do we mean by gender identity? Well, gender identity is something that changes payoffs from different action as dictated by some prescriptive norms. What might those norms be? Well, one norm would be men should not do women's work in the house, in the home, like cooking, cleaning, and so on. Some men might consider that as women's work. And they might feel particularly high disutility from doing so if they violate that law.

Another norm would be men should earn more than their wives. Now husbands now lose identity, experience lower utility in circumstances when these prescriptions are violated, when he does housework, or when his wife earns more than half of the household income. Notice that wives might also lose identity. For simplicity, we're going to talk more-- only about husbands, I'm doing so for now.

Now Bertrand et. al. particularly focus on the prescription that men should earn more than their wives. And it's important to notice that this particular gender identity norm would only matter in a world where women could-- would not matter in a world where women could never earn more than her actual or potential husband. But if women are held back by technological factors, by education, et cetera, and never even earn close to as much as their husbands or potential husbands, well, then the gender identity norm is kind of irrelevant.

However, as women make gains in the labor market, these slow moving gender identity norms can become increasingly relevant and increasingly relevant constraints. So now as women have more and more earnings potential and are much-- be quite likely to earn more than their husbands, these norms really bite and might hold women back.

Now how would you study this question empirically? The idea here in Bertrand et. al. is that, well, if husbands lose identity when their wives earn more than half the household income, well, then they will try to avoid such situations, either-- in various ways, and particularly by avoiding such marriages in the first place. But also if those marriages happen, that might lead to divorce and so on and those marriages might break up.

So the idea of a paper then is to look for a missing mass in couples where, particularly in places where the husband earns just a little bit more than their wife. Now how do they do that? Well, they have to look at the distribution of relative income using US admin data. What do I mean by relative income? Relative income is the income between the husband and the wife.

They have essentially, there's a very nice survey, data of the survey income and program participation. These are a series of representative national panels that about 70,000 couple observations from 1992 to 2004. This includes couples only in which both the husbands and wives are earning positive income. So notice that this was not about husband and wife working at all. It's not about women being at home versus working. But it's about cases where both couples or both spouses have positive income.

Their income measure is annual total labor income plus self-employed income as well. Now using that data, Bertrand et. al. can now compute the shares of couples earning different fractions of total income. So what I'm going to show you is the fraction of the wife as a share of total income in the household. And that's graphed using 20 bits.

What does this graph look like? You can see here the share earned by the wife is just a fraction of couples. This overall. So this is all the couples. This is [INAUDIBLE]. This all adds up to 100% overall.

And what you see is, perhaps not surprisingly, that the share-- the fraction of couple for which the share of earned by the wife is higher than 50% is not super high. That's not surprising given that men on average are earning more than women. So that's not the question here. The question here is, is there a missing mass just above 50%?

The idea is, suppose there are some couples where women earn 45% or 46% or even 49%, there the norm does not cause an issue or there's no identity problems because the husband earns somewhat more and feels happy about it. But as soon as you cross the threshold to like 51%, now there's a problem. Now the woman earns more. The husband feels uncomfortable, unhappy and so on.

And that might lead to those marriages disappearing in various ways. A, they might not form in the first place. Potential husband and wife might never even marry because husband gets uncomfortable or does not want to marry somebody who earns more. Second, there might be divorces. These marriages might essentially dissolve because there's lots of conflict in the household.

Now, how does that manifest here? You can look at like if you come-- if you look at the graph and look at this line here coming from the left was 50%, you can kind of look at like, try to predict-- and it's kind of the idea of the empirical strategy-- you can try to predict what fraction would you expect there to be just above 50%?

And when you do that, you think, you know, the fraction should be about something like 7%, should be about here. But, instead, there's this jump down to about 6%. So there's about 1% of couples are missing here. There's a missing mass going on here, which suggests that there are some marriages that are do not happen or dissolve because of those identity concerns.

This is what I already said here. There's a cliff to the right of 0.5 in the distribution of relative earnings in most couples. And that's an implication of the prescription that men should earn more than their wives.

Now what are the mechanisms behind this missing mass? Bertrand et. al. had a wealth of evidence, additional evidence showing these mechanisms. In particular, they show three mechanisms, which is, first, missing couples did not form in the first place. When you look at-- when you just randomly match men and women, some of which were men have higher earnings potential compared to women and the other way around, the couples where women are earning just somewhat less than men are just much less likely or less likely to marry compared to couples where women are just earning somewhat less than men.

Second, such couples are less happy and stable and more likely to end in divorce. So conditional on marrying, they're more likely to get divorced as marriages just end unhappily. And that's, of course, then also contributes to this missing mass.

And then third, in cases where wives have higher earnings potential than their husbands, they decide now to work less outside of the household or, you know, put in less [INAUDIBLE] work so they get less promoted, which then sort of pushes down their earnings or increased husband's earnings in relative terms, such as done in such cases the woman earns less than 50%. And that also leads to that missing mass. In fact, what you see here is this increase in mass going from like 20% to 40%, which perhaps is also like the reason because some of the mass that's missing here is pushed here to the left.

Now, in addition to working less outside of the household and just working fewer hours and working less overtime, it's on, women then are also held back by non-marketing in childcare work. So women, in particular, women who are earning just a little bit more than their husbands are doing more nonmarket and childcare work, which is often called the second shift. So the idea here is that husbands even when his identity gets threatened by the woman earning more than he does, and then he can particularly not to the childcare or nonmarket work. So the woman ends up not only working more outside of the household and earning more and being more productive in that work, but also doing more work at home and in addition to that. That's essentially what's called the second or double shift.

Now an additional issue that is a really important contributor to the gender gap is the arrival of children. And this is a very nice paper by Kleven et. al. that looks at people's earnings, in particular, the earnings, how earnings devolve after the birth of a couple's first child. This is an event study that looks at years over time for men and women.

And what you see essentially for men, these are earnings and on the right here you have hours worked. For men, essentially, there's barely any difference, if any difference, of earnings once the first child is born. Essentially, men's earnings tend to be pretty flat. In contrast, of course, that could be for the growth that's avoided. But, you know, at least there's no reduction in men's [INAUDIBLE]. This is, by the way, Danish data.

Now for women, however, there's a clearer reduction in earnings. And the long-run child penalty is about almost 20%. So, essentially, women after-- if you look at men and women that earn the same to start with, women earn about 20% less in the long run. That is due to the first child or after-- this gap arises after the first child was born.

This is true for earnings. It's also true for hours worked. And there the penalty is about 10%. It's also true for wages, as in how much are people paid conditional on work. So part of that is women working less. But part of it is also women earning less, a la being promoted less, being paid less, getting fewer raises and so on once the child is born.

One very interesting fact in this paper is that child penalties are transmitted through generations from parents to daughters, suggesting an influence of child environment on gender identity. There seems to be something about the gender identity about the woman is supposed to take care of the child and that runs in the family. And that leads to particularly large what they call child penalties for those kinds of women compared to men.

Now one thing you might say is, well, couldn't we have some gender neutral family policies that alleviate that? And, in particular, in academia, for instance, one idea is one could use gender neutral tenure clock stopping policies to alleviate such issues. And the idea is that, in particular, sort of early years of the career are particularly valuable.

So the very least we can do is we can stop the tenure clocks to help women or couples to deal with childcare. And many, many universities and research intensive universities in the US have adopted such what's called gender neutral tenure clock policies in the US.

Now what do I mean by gender neutral? Essentially, if a couple has a child, not only does the woman, if she has given birth to the child, does she get an extension in the tenure clock, but also the spouse will get such an extension as well. So it's gender neutral in that sense.

Now, what would such policies then do? Well, they're intended to involve men more in child care and foster that, the idea being like if you only give that extension to women, it sort of pushes women to take care of their children and sort of not to work. And perhaps if we involve both spouses or parents in the childcare, providing those gender neutral tenure clock stopping policies might contribute to that.

But, of course, there's no enforcement and thus, potentially, such policies in that potentially these policies might be even enhancing gender inequality. And, particularly, if it's the case that if men take advantage of this tenure clock stopping policy by essentially just having another year to do research, but still don't do very much in supporting the child, then men might actually benefit from it, while women might not benefit at all in part because, relatively, they're going to do worse compared to men.

And so Antecol look, in fact, at the impact of these adoptions of these supposedly gender neutral tenure clock stopping policies. And they find that the introduction of this policy at [INAUDIBLE] economics department substantially reduced female tenure rates while increasing male tenure rates. So essentially what looks like a neutral policy, in fact, is not gender neutral at all. It makes things worse for women and better for men.

Let me tell you about one final paper on gender identity and the labor market. This paper is called "'Acting wife'-- marriage market incentives and labor market investments." It's by Bursztyn et. al. and it asks the question whether women avoid career-enhancing actions because these actions signal undesirable or seemingly undesirable traits such as ambition to the marriage market.

There's two parts in the paper. One part is just observational data that finds that while married and unmarried female MBA students perform similarly when their performance is unobserved by classmates, such as on some exams or problem sets, unmarried women have lower participation grades which are observed by their classmates. And so the idea here is that only unmarried female MBA students have incentives on the marriage market.

And, therefore, when their participation is observed, when their performance is observed, they might want to scale back their ambition because they look too ambitious on the marriage market. And that might be adversely interpreted by potential dates. And, of course, married women, since they are already married, don't have that incentive. So they don't engage in such behavior.

As you might know, there's lots of dating going on. So lots of MBA students are actively looking for a spouse during their experience, both men and women. Now in addition, they have field experiments with MBAs in which they vary the expectation whether responses in a real stake placement questionnaire is going to be made public, in particular, whether it's going to be observed by their peers. And now female, single female students reported or desired salary and willingness to travel and work long hours on those questionnaires when they expected their classmates to see their preferences.

That is to say, when these questions are public, single female students reported lower ambition as manifested by desired salaries and willingness to travel and work hours, presumably because they want to look more favorable on the marriage market. Notably, other groups' responses were affected by such peer observability. So in other groups, it didn't matter whether it's in public and in private.

In addition, there's a second experiment that indicates the effects were driven by the observability by a single male peers. That is to say, if another woman saw the information, that didn't matter so much. But if this other single male peer was likely or possibly going to see that information, then these effects would show up. So that's overall showing that, essentially, what's perceived to be undesirable traits, such as ambition, if that is perceived to be damaging on the marriage market, that can then, in turn, reduce ambition and then contributes to the gender gap overall.

OK, let me tell you now about a final paper on saying no and on demand and supply for different tasks. This is a very nice paper by Vesterlund et. al. And it asks the question whether women say no often enough. And so that's motivated by the observation that female faculty member spend fewer hours in research and more hours on university survey committees than male faculty. They are also more likely to have positions on university-wide committees. They advise more undergrad students and participate more in departments and college level committees than male faculty.

More generally, in mid-level jobs, men more than women evaluate the individual task assignments as challenging. And this is partially attributed to differential task assignments by supervisors. So if you sort of see these differences, you might ask the question, why do women decide to spend or spends their work time differently?

And there's two types of dimensions here. One is demand and one is supply. So demand is essentially the question whether sex differences in the types of tasks that women and men are asked to do at work is that [INAUDIBLE] essentially what they're being asked. And are women more likely than men to be asked to do what's called non-promotable tasks, tasks that are not really helpful for their career, for getting promotions. And then you might ask why that's the case.

Supply is are there sex differences in the willingness to agree to perform non-promotable tasks when asked? [INAUDIBLE] conditional on being asked, were women more likely to say yes. And, again, are women more likely to demand to say yes to non-promotable tasks.

Now why do we care? Well, there's an individual decision-making perspective, which is that people might make suboptimal decisions how to allocate their work-- time at work. And understanding these underlying reasons and [INAUDIBLE] understand these underlying reasons, which will lead them potentially to some interventions to improve decision-making for certain individuals, which have a lot of what behavioral economics is. But there's also managerial and social planner perspectives that organizations may not be using their resources most effectively. And if you sort of reallocated some of the tasks, that would increase output overall.

And then finally, there is a public policy perspective. Well, if these sex differences in the allocation of time explain vertical sex segregation, essentially women not being promoted enough, then it would help us sort of improve gender equity overall and perhaps also try to reduce the gender gap. Now what are these promotable and non-promotable tasks in the [INAUDIBLE] academics? A promotable task is a task that's doing research. That's essentially what people are evaluated on eventually, but also any other task, essentially, that's seen where other people will reward you for it when you do it.

And non-promotable tasks are essentially tasks that often [INAUDIBLE], many people could do the task and everybody wants the task to be done, yet everybody prefers somebody else to do it. An example would be an ethics committee of the university. Surely, we should have one. But surely, sitting on the committee will not help people get promoted or get tenure when it comes to research.

Now the authors had to do a field study of faculty at a large public university. And they sent, essentially, emails from the chair of the faculty senate asking them to volunteer to join one of several university-wide faculty senate committees. And the clear answer in this study is that women are much more likely to volunteer when asked.

Now, that is an interesting fact. But it doesn't necessarily let us disentangle the demand and supply explanation that I showed you previously. So in addition then, the authors do a lot of experiments in which they do what's called the threshold public good scheme. And this is sort of very much trying to mirror or resemble the reality. In this game, a small group that needs to find a volunteer for a task.

And participants are anonymously matched into groups of three. They're randomly every match for each of 10 rounds. The game is set up such that everyone prefers that the task be undertaken by someone other than themselves. People get two minutes to decide whether to invest. And so the task is essentially you have to click a button. And then you invest the money. Only one person can invest. And the round ends when somebody invests.

If no one invests, group members all earn \$1. If a person invests, that person earns \$1.25 and the remaining group members earn \$2. The clock ticks down until one person invests or no investment is made in two minutes.



So the game is essentially set up in a way such that everybody wants that somebody-- and it's clear that somebody should invest. But nobody wants to invest because if you invest, you only \$1.25. But if you don't invest or somebody else invests and you don't invest, you get \$2.

Now in this game, women are significantly more likely to invest. We can essentially see the probability of investing here on the left side by round. And the red line, which is the upper line, the woman line, the female line is way, way higher compared to the male line. So there's much higher probability of investing in this game.

Now why are women more likely to invest? Is it the case that women believe that their cooperation is necessary for an optimal group decision, but men believe that their cooperation isn't required. So it could be that women think that, well, somebody-- if I don't do it, nobody else does it. And then men essentially think, well, somebody else can do it, in particular, women could do it if they're in the group.

So then the authors do the same experiment again using single sex sessions where essentially all three people in the session are female or all people in the session are male. And then what they find is essentially, interestingly, these differences in gender now go away entirely. There's no gender differences in same sex sessions anymore.

So just to summarize, so experiment one, when there's a mixed gender, women are more likely to invest. And experiment two, women and men are equally likely to invest. Now, one question you might ask is, well, is it the case that really, isn't this really about beliefs?

So it could be that there's no, in fact, no differences in the preferences for investing. So it's not like women are necessarily nicer than men. But, rather, there are differences in beliefs that the women will invest. In particular, if there's a woman in the group, everybody else will think, oh, the woman is going to invest.

And the woman might think like, well, if others don't invest or others will not invest, and she might believe that, she might have to or think that she needs to do that herself, otherwise everybody will end up with \$1 and less money.

So now there's a third experiment here where people are asked to essentially pick whom they would like to ask. And so there's four people per group. And three people can invest. And the incentives are [INAUDIBLE]. These are like-- they're called the green players. And one person is unable to invest, but asks one of the three to invest. Importantly, the requests are not even binding. That's the red player. The red player, essentially, once the [INAUDIBLE] happens, you incentivize that that happens.

So this is kind of what this looks like. You would get like three different players here. And you can essentially decide who you would like to trust. And it's set up, of course, that you can see the gender. Some are male and some are female.

Now women are way more likely to be asked. This is the total times asked to invest for male and female players. This is the relative frequency. And essentially, you can see their distribution here is way shifted to the right for women compared to men.

Now is it better to ask a woman? Well, absent a request, the investment rate actually does not differ by gender. But when asked to invest, women are more likely to comply. So when the woman is asked to invest, 76% of women invest compared to 14% when not asked.

In contrast, men are also more likely to invest when asked, but it's only 51% compared to 14% when not asked. So the marginal increase of being asked is higher for women than for men. So if you're thinking about, like whom should we ask? Well, you kind of want to ask a woman because a woman is more likely to comply with your request.

And so now if women-- since women are more likely to be expected to say yes, they also ask more. So in that sort of sense, you know, the gender differences get amplified by increased demand for women to contribute. Well, particularly, if there's a man and a woman where both can think about like one of them could do it, the man will think, well, the woman will do it anyway.

The woman will think, well, the man is not going to do it. And he's going to think that I will do it anyway. And so then the woman ends up doing it more.

In addition, when there's an opportunity to ask somebody, but asking is kind of costly and you just want to find somebody who will do it. But who will you ask? You ask the woman. Because she's more likely to comply, more likely to say yes.

Let me sort of summarize what we discussed. So first is large gender, wage, and earnings gaps. They have been reduced due to technological advances and other improvements over time. But there are still persistent gender differences in the US and many other countries.

Bias beliefs and identity concerns play a major role in explaining these differences. In addition, there's some feedback mechanisms between the demand and supply of non-promotable tasks that could be, potentially quite important. Of course, I showed you only a lab experiment. but these could be really important in real-world situations. So not only is it that what matters is what women decide to do, but also what they're asked to do by others, in part as a response to their propensity to say yes or no to certain tasks.

Better understanding these issues can help us mitigate the gender gap. And we can sort of then, by understanding these issues better and particular beliefs, you might be able to improve or correct biased beliefs and, therefore, close the gender gap potential.

What's next in the next few lectures, lecture 19, we'll talk about frames, defaults, nudges, and mental accounting. Please read Madrian and Shea 2001. And lecture 20, we'll look about-- talk about malleability and inaccessibility of preferences. Please read Ariely 2003 for that. That was all I have to say. Thank you so much.