Lecture 18: Alternative models of discrimination.

Becker's models are in some ways an argument for how markets can eliminate the effects of prejudice on workers' wages. He starts from the assumption that prejudice is a facet of psychology or preferences that's separate from markets, or precedes markets, and doesn't play any role in the actual process of production. He then asks whether it can affect market wages in various circumstances.

These models take a different approach. Instead of seeing prejudice as separate from the production process, they think about the way that prejudice influences the production process, and in turn the way that the production process influences prejudices. In other words, they ask when market interactions *perpetuate* prejudices rather than eliminating them.

We'll be talking about three models today:

1) Statistical Discrimination

2) Coate & Loury—discrimination and human capital investment

3) Basu—discrimination as a focal point

Model 1: Statistical Discrimination

This is a concept that we will talk about more in later lectures, but it's important to understand this idea before we can really understand the Coate and Loury model.

The idea behind statistical discrimination is: it may be in someone's best interest to discriminate against a certain group (women, African americans, etc) because membership in the group is *associated with* some characteristic that's important to me.

So, for example, let's say that you're recruiting salespeople for a yacht company, and you're trying to decide what wage to offer each person. Your goal is to offer someone a wage no higher than the expected revenue they'll bring in by selling a yacht. You know each applicant's grades, previous experience, etc, but you don't know if they have any rich friends who might be interested in buying a yacht (and everyone interviewing for a yacht salesperson job will name drop a bunch of rich people, so you can't really tell).

However, you know that, on average, white applicants are more likely to have wealthy friends than are black applicants. To make this concrete, let's say that ¾ of white applicants have a wealthy friend, and ¼ of black applicants have a wealthy friend (not real numbers!).

Further, let's say that whereas the likelihood of making a sale is 4/5 if you have a wealthy friend, it is only 1/5 if you don't (for simplicity, you either sell a yacht or you don't). Note that race doesn't matter for your productivity at all, other than as an indication of your likelihood of having a wealthy friend.

So, we can determine the likelihood that a white or black employee will sell a yacht as:

	Frac. Rich Friend (4/5 chance of sale)	Frac. No Rich Friend (1/5 chance of sale)	E(Yacht)
White	¾	¼	¾ *4/5 + ¼*1/5=13/20
Black	¼	¾	¼ *4/5 + ¾ *1/5 = 7/20

And, if we suppose that selling a yacht generates $200,000 for the company, we'd expect the company to offer potential white employees $130,000 starting salaries, but only offer black employees $70,000 starting salaries.

In this case, are employers discriminating because they're *prejudiced?* In other words, are they discriminating based on negative perceptions, tastes, or opinions of one group of people compared to another? Kind of—employers are offering lower wages to black workers because they perceive black workers to be less likely than white workers to have rich friends who might want to buy a yacht.

The problem is, in this case, that negative perception is correct, and based on evidence. Whereas employers who engaged in what we called "taste-based" discrimination against minority workers suffered lower profits as a result, and would eventually be driven out of competitive markets, employers who engage in statistical discrimination will earn higher profits than those who don't. A firm that pays white and black yacht salespeople the same amount will earn lower profits than one that pays differently, and indeed will be driven out of the market in a competitive equilibrium.

Another way to think about this is to ask: what if instead of having a human decide salary offers, we had a computer program decide, where the program used all of the available information to predict a salesperson's success. The computer program isn't prejudiced in the way we normally understand the term—it treats the variable "salesperson race" in its dataset the same way it treats any other variables, and has no preconceived notions about it. But given the scenario we're discussing, the program would end up offering black workers less money than white workers, because it would accurately predict that black workers would be less productive.

Is statistical discrimination *really discrimination?*

Part of the reason we care about discrimination is because we think it's morally wrong to judge people on characteristics like race, or sex, or immigration status. Using race, gender, sex to guess about someone's behavior, attributes, productivity, etc, *is* a way of judging people based on those characteristics, and it *is* discrimination. Even if these guesses are right *on average*, they're not right across the board, and they make it more difficult for people to exist in the world as individuals. We'll be talking more about this through this lecture.

However, it's impossible, from the perspective of a researcher, to tell whether employers are engaged in statistical discrimination or if they have access to information about individuals that a researcher lacks. For instance, in the model above, employers would offer an average wage of $130,000 to white employees and $70,000 to black employees even if they did know who did and did not have rich friends.

Model 2: Coate and Loury: Prejudice as a Self-Fulfilling Prophecy

The second model that we're going to talk about takes the idea of statistical discrimination and extends it, by thinking about how *being statistically discriminated against* affects the choices of white and black workers (or any more and less advantaged groups). The model asks: If there are no intrinsic differences between two groups of people, can negative perceptions of one group become a self-fulfilling prophecy? That is, can the belief that black workers are less productive on average than white workers, for example, reduce the benefit of investing in skills for black workers until they are indeed less productive on average than white workers?

Let's think through this with a slight simplification of the Coate and Loury model.

We'll say that there are two types of jobs: laborers and engineers. Because engineers are paid more highly than laborers, and because (let's say) it'll take a year for employers to see whether a worker was productive or not, all workers would rather be engineers than laborers, regardless of how productive they'd be in the job. However, workers will only be productive as engineers if they've made a costly investment in skills and training. As a result, employers need to determine who's qualified to be an engineer by maintaining some hiring standard (say, giving everyone a competency test). If, after reviewing the test and any other information, they think that a worker is likely enough to be qualified for the engineer job that their expected profit from hiring them is positive, they'll hire them.

The Firm:

To make this more concrete, let's say that a qualified engineer produces $200,000/year for the firm, and an unqualified engineer produces $0/year for the firm. The firm pays engineers $100,000, so they make $100,000/year in profit if they hire a good engineer, and -$100,000 if they hire a bad engineer.

As a result, the firm's expected profit from hiring a engineer is:

$$E(\pi) = P(qualified) * 100{,}000 + \big(1 - P(qualified)\big) * -100{,}000$$

And they will hire a engineer as long as that expected profit is greater than zero:

$$\begin{aligned}
&P(qualified) * 100{,}000 + \big(1 - P(qualified)\big) * -100{,}000 \geq 0 \\
&P(qualified) * 100{,}000 * 2 - 100{,}000 \geq 0 \\
&P(qualified) \geq \frac{1}{2}
\end{aligned}$$

So they're hire a engineer as long as they think it's more likely than not that the engineer is qualified.

To find out whether a worker is qualified (S=1), they'll give the worker a test. The worker will definitely pass the test (Z=1) if they're qualified, but **1/3 of unqualified workers also pass the test.**

The Workers:

Workers in this model (at least, our simplified version of it) are all identically capable to begin with, and all start off unqualified to be engineers. However, if they make a human capital investment that costs them $20,000, they'll become qualified (S=1).

Workers need to decide whether that investment is worthwhile.

Let's say that if they aren't offered a engineer job, their alternative wage is equal to $40,000.

Let's say that they know that they'll be hired if they pass the test, but won't be hired as a engineer if they don't. In that case, if they study to become qualified, they'll earn:

Earn(S=1): $100,000 - $20,000 = $80,000

Earn(S=0): $40,000*2/3 + $100,000*(1/3) = $60,000

So, if they know that they'll be hired if they pass, it'll be worth it for workers to study, and all workers will study.

The Firm:

So, given this, will a firm want to use the rule we just talked about? Will they be willing to hire workers who pass their test? Yes!

What we said before is that the firm will hire a worker as long as they think the likelihood that that worker is qualified is greater than ½.

So what's the likelihood that a worker who passed the test is qualified? Remember that there are two ways to pass the test:

1) Be qualified

2) be unqualified but pass the test by luck.

So the likelihood that someone who passed the test is actually qualified is just the relative number of qualified people compared to lucky unqualified people. Specifically:

$$P(S = 1|Z = 1) = \frac{Qualified\ people}{Qualified\ People + Lucky\ Unqualified\ People}$$

$$P(S = 1|Z = 1) = \frac{1 * P(S = 1)}{1 * P(S = 1) + \frac{1}{3} * \left(1 - P(S = 1)\right)}$$

If they know that everyone will want to get qualified, this probability is 1, because P(S=1)=1—everyone's qualified. So we have a happy, simple, straightforward solution to our little problem—a passing test score gets you a job as a engineer, everyone gets qualified, and everyone passes the test. Yay.

Race:

So far, our happy model hasn't distinguished people at all by race—everyone is identical. But now let's say that some workers are Black and some workers are White. Let's further say that there's no actual difference, whatsoever, between black and white workers. No difference in the number of rich friends, no difference in the costs of getting qualified, nothing.

However, firms *believe* that black workers are different from white workers. In particular, firms believe that at **most 1/5 of black workers can possibly be qualified** to be a engineer, because of racism. Essentially, firms have wrong, outdated, negative stereotypes about black workers based on no real differences.

Will these stereotypes affect the market? Will the be eliminated by competitive pressures? Sadly, no.

To see this, let's ask the question of whether a firm that believes that no more than 1/5 of black workers are qualified will hire a black worker who passes the test.

Remember, they'll hire the worker if they believe that they've got a 1 in 2 chance of being qualified:

Hire if:

$$P(S=1|Z=1) = \frac{1 * P(S=1|B)}{1 * P(S=1|B) + \frac{1}{3} * \left(1 - P(S=1|B)\right)} \geq \frac{1}{2}$$

But now, instead of P(S=1)=1, we have P(S=1|B)=1/5

So:

$$\frac{1 * \frac{1}{5}}{1 * \frac{1}{5} + \frac{1}{3} * \left(1 - \frac{1}{5}\right)} \geq \frac{1}{2}$$

$$\frac{\frac{1}{5}}{\frac{1}{5} + \frac{4}{15}} \geq \frac{1}{2}$$

$$\frac{3}{3+4} = \frac{3}{7} \geq \frac{1}{2}$$

So, since the firm thinks a black worker who passed their test can't have more than a 3/7 chance of actually being qualified, they won't hire black workers as engineers, regardless of their scores.

Decision of Black workers:

If black workers know that they won't be hired regardless of their test scores, will any black workers decide to invest $20,000 in getting qualified? No! Waste of money! And as a result, every single black worker who passes the test will do so by luck. In reality, P(S=1|B)=0. What has happened is—what started as taste-based discrimination turned into statistical discrimination.

Stability of the Equilibrium:

The tragic conclusion of this model is: even though black workers are no different from white workers, the existence of prejudice means that black workers make different choices than white workers. Those choices, in turn, justify the initial prejudice against black workers, and will result in discrimination *even if everyone understands what's really going on perfectly*.

What do I mean by this? Imagine that you were able to explain this whole model to an employer, and convince them that they were living in it. They became absolutely certain that black and white workers were intrinsically identical. Would it be in their interest to hire black workers who passed the test? No, because since no black workers have an incentive to study (since no one else will hire them), all of their black applicants will have passed by luck.

So, rather than working to eliminate negative stereotypes, market forces in this case work to *preserve* discrimination, even if everyone was made to understand that differences between groups were generated entirely by that discrimination. Everyone in this model becomes trapped in a wasteful and unfair equilibrium, and no one has the power to change it.

Affirmative Action:

How could you change the equilibrium in this model? By legally requiring companies to have equal standards for black and white workers. If you told companies that they had to use the same test to determine black and white workers' eligibility, the result would be that employers would hire black workers who passed the test, even if they thought it was luck. This would then create the same incentives for black workers to invest as for white workers, and we'd end up in an equilibrium where everyone had the high-wage job and everyone was qualified.

You could also get here with a quota system: If you said that employers had to hire the same fraction of black applicants as white applicants, the natural place for employers to start would be to hire black applicants who passed the test (initially 1/3 of black applicants, rather than 100%). This would again create an incentive to pass the test, and thus an incentive to study, and we'd escape the bad equilibrium.

Note: Escaping the discriminatory equilibrium is good for everyone in this model. Essentially, it allows black workers to invest productively in their skills, which increases the total talent available in the economy and allocates talented workers to the jobs where they're most productive. A working paper by Chang-Tai Hsieh and Erik Hurst (The Allocation of Talent and U.S. Economic Growth, Hsieh, Hurst, Jones, and Klenow) estimates that somewhere between 20 to 40% of the growth in the US economy from 1960 to 2010 can be explained by the integration of women and ethnic minorities into skilled professions. Women and minorities have gone from making up 6% of doctors and lawyers in the US in 1960 to making up 38% of doctors and lawyers by 2010, and are an even larger share of new doctors and lawyers.

Model 3: Discrimination as a Focal Point

The final model is mathematically very simple, but makes a very deep and challenging point about discrimination. Up until now, we've thought about prejudices as originating outside market interactions. Depending on the model, markets can either allow those prejudices to metastasize into discrimination, or can sweep these prejudices away. In this model, Basu considers the possibility that markets can actually generate discrimination as a tool to increase productivity. In this argument, when ordinary people engage in economic interactions, looking for strategies that can help them improve their earnings, they will in some cases settle on discrimination as a way to do so, even in the absence of *any* real differences between groups.

How can this happen? Consider situations where it is in your interest to work with, support, or partner with someone that you expect to also get the support and partnership of others. For instance, let's say that your friend's band is creating their first release, and asks you to direct their music video. They offer that if you do it, they'll give you a share of the proceeds. Regardless of what you think of your friends' music, you might be more likely to do it if you think that your friends' music is likely to be to the taste of excellent producers and promoters, if you think that your friend can get talented and successful musicians to provide guest verses, or if you think that big venues are likely to book your friends when they go on tour. In other words, a big part of whether it's smart for you to invest in your friend comes down to whether you think other people are likely to invest in your friend.

And, importantly, those other people *also* only want to invest if your friend's band can get you on board. The venues are going to be more willing to book a new band if they think it can attract a great music video director, and a great producer, and great guest verses. Other artists are more likely to supply guest verses if they think that your friends' music will be successful, part of which boils down to the music videos, venues, and producers. In other words, you and all of these other players are playing a *coordination game*, where your investment is most likely to pay off if others invest in the same person.

What does this have to do with discrimination? Basu shows that, given the very simple assumption that you are best off investing in someone that others also invest in, society can easily (and almost inevitably) fall into equilibria with favored groups and disfavored groups.

To see this, imagine that there are four competing startups, each of which builds network infrastructure for companies. A large local company is deciding which startup to hire for its services, and an investor is deciding which company to invest in. If the investor chooses the start-up that receives the contract, the investor is more likely to get paid back because of the fantastic revenue stream. Meanwhile, if the company chooses the start-up that gets the investment, it will be more likely to get reliable, high-quality service because that start-up will be more capable of scaling up effectively and paying for high-quality service.

To make things simple, we'll say that the payoff for both companies of picking the same startup is 2, and the payoff for picking different companies is 1.

Now let's say that two start-up founders are men, and two are women. Gender plays no role whatsoever in the start-up's operations—they're identical in all ways except the gender of their founders. However, because the founders' genders are different, the investor and the buyer each choose whether to intentionally select a male founder, to intentionally select a female founder, or to be gender-neutral.

	Buyer			
Investor		Neutral	Male	Female
	Neutral	5/4 2*1/4+1*3/4 =5/4	5/4	5/4
	Male	5/4 5/4	2*1/2 +1*1/2 = 3/2 3/2	1 1
	Female	5/4 5/4	1 1	3/2 2*1/2+1*1/2 =3/2

Essentially, both players can get some modest payout of 5/4 on average if they don't discriminate by gender—that is, if at least one player gives men and women equal opportunities to win the investment. But if both players discriminate, and discriminate in favor of the same group, their likelihood of choosing the same startup increases, and with it their payout.

So, while it's possible that no one will choose to discriminate, once someone starts discriminating, it's always best for others to match them. What has happened is that there are benefits to both sides to shrinking the field of consideration, in order to make coordination easier. If it's most productive for all relevant actors to pick the same start-up, racial discrimination, or gender discrimination, or college prestige, or any number of other high and low-status group memberships can be a way to accomplish that.

What this might look like in practice (especially when these dynamics are hidden), is that once a story starts spreading about one group of people being particularly well-suited to a particular line of work, that reputation becomes a self-fulfilling prophecy. The favored group will **in fact** become more productive, not because they've changed but because they're benefiting from increased investment, cooperation, and engagement with several different facets of their industry or their work. And that means that the story about one group being *intrinsically better* than another will feel true to more people and get stuck.

And of course, these models interact with each other. If you are a member of a group favored in some profession, it will be more valuable for you to invest in the skills needed in that profession than it would be for someone from a disfavored group, making you yet more productive. This means that once a discriminatory equilibrium is set, it can be very hard to dislodge.

It also means that, in contrast to the Coate and Loury model, affirmative action and other policies aimed at eliminating discrimination won't be a one-time fix. In Coate and Loury's model, once affirmative action is in place for a short time, everyone realizes that white and black workers are the same and stops discriminating. In this model, if someone required that half of a firm's investments were women-owned businesses, they would succeed in breaking out of a coordinating equilibrium only as long as the policy was in place. Once it was released, there would be a drift back toward discrimination, because discrimination is productive.