1. Introduction

In the past three lectures, we explored three rival theories of equality of opportunity—meritocracy, fair equality of opportunity, and luck egalitarianism. And we then took a more practical turn, by examining discrimination and discriminatory hiring practices in light of these accounts of equality of opportunity. This last lecture will examine the ethical standing of another practice: affirmative action, or preferential hiring. Affirmative action, by contrast with discrimination, is usually taken to be a good thing by political philosophers. So one of the questions we'll be interested in here is whether or not the theories of equality of opportunity we've canvassed can provide support for this.

2. Defining Affirmative Action

Roughly, affirmative action refers to a range of policies that give preference to members of disadvantaged or marginalised socially salient groups on the basis of their membership in that group. Often, affirmative action is discussed in the context of hiring for jobs, of selecting candidates for social positions, or of university admissions. It might take the form of a quota system whereby a certain percentage of jobs or vacancies must be reserved for members of a certain racial minority. Or it could involve giving ‘extra points’ to applicants from certain groups. Affirmative action sometimes goes by other names. Thomson for instance uses the term “preferential hiring”, while others talk about positive discrimination. The legal standing of affirmative action differs across different countries. In France, it is legal. For instance, a 2011 law required firms to ensure that 40 percent of board members in all stock-exchange listed or state-owned companies be women. That would be an example of a quota approach to affirmative action. In the case of the US, the Supreme Court declared in 2003 that affirmative action by universities was legal (Grutter v. Bollinger). However, the Supreme Court also said that using racial or gender quotas was unconstitutional. So the idea is that it is legal to take race or gender into account when choosing between applicants, but that employers shouldn’t have strict numerical targets of the kind implemented in France. In the UK, affirmative action is generally considered to be against the law. The reason for this is that affirmative action allows employers to appeal to socially salient categories like race or gender to prefer one candidate to another. And this, at least at first sight, seems to violate the provisions against discrimination put forward in the Equality Act of 2010 (see lecture 3). Now, do note that there are exceptions to this general prohibition (for an overview, see https://en.wikipedia.org/wiki/Affirmative_action#United_Kingdom). In sum, there is substantial political dispute about whether affirmative action is a good or bad thing. This constitutes a political reason for exploring its ethical standing.

But there’s another reason for exploring the ethical standing of affirmative action. While there is a lot of disagreement in politics about whether affirmative action is a good or bad thing, political philosophers on the whole tend to support it—they mainly disagree about what exactly justifies affirmative action. But this might seem strange, because affirmative action seems discriminatory. Recall from lecture 3: X discriminates against Y when (1) X
treats Y differently from Z; (2) the difference in treatment is believed to be disadvantageous to Y; (3) the difference in treatment is due to Y and Z being from different socially salient groups. Affirmative action seems to satisfy these conditions: preferential hiring for, say, people of colour involves treating them differently, on the basis of their race, in a way that appears to be disadvantageous to other candidates.

This generates a puzzle:
(1) most political philosophers think that affirmative action is a good thing;
(2) affirmative action seems to be a form of discrimination;
(3) most political philosophers think discrimination is typically wrong.

Are political philosophers simply confused? To shed light on this puzzle, what we need to do is look more closely at the moral foundations of affirmative action. What, if anything, justifies it? Can the moral justification for affirmative action be reconciled with thinking that discrimination is most often wrong? When answering these questions, we'll be especially interested in considering affirmative action through the lenses of equality of opportunity. In other words, we're going to consider whether or not equality of opportunity can lend support to affirmative action; if it does support affirmative action, one thing we need to find out is whether it can do so in a way that preserves the intuition that discrimination is very often wrong.

3. Luck egalitarianism and affirmative action

We've seen several theories of equality of opportunity. Here, let's start with luck egalitarianism. Luck egalitarianism, recall, states that we should eliminate inequalities in well-being if and only if they result from bad luck rather than from genuine choice. Now, last week we saw that luck egalitarianism does not entail that discrimination is necessarily wrong. For luck egalitarianism, treating some groups better than other groups on the basis of their race or gender is wrong only when it undeservedly creates inequality in opportunities for welfare between these groups. And, at least in principle, it seems possible that some instances of discriminatory hiring overall do not generate inequalities in opportunities for welfare. For instance, this might be the case if the person being discriminated against started out with far more opportunities for welfare than anyone else. The fact that luck egalitarianism doesn't necessarily oppose discrimination is important in our context: it means that even if affirmative action is discriminatory, luck egalitarianism doesn't necessarily oppose affirmative action.

In fact, Shlomi Segall (2012, ‘What’s So Bad About Discrimination?, p. 96) suggests that luck egalitarianism can actually support affirmative action. The reason is that affirmative action gives preference in hiring to people who come from relatively disadvantaged social groups. So, for Segall, since these groups currently have fewer opportunities for welfare than other groups, giving them preferential treatment in hiring may be a way of creating more equal opportunities for welfare. For example, giving preference in hiring to black Americans may be a way of reducing the vast racial inequalities that persist in America.

In this light, it seems as though luck egalitarianism might help with our puzzle.
- On the one hand, luck egalitarianism helps capture the intuition that affirmative action is morally justified—an intuition shared by many political philosophers and politicians.
- On the other hand, luck egalitarianism seems capable of capturing the intuition that many instances of discrimination are wrong—as, for instance, when black Americans weren’t allowed to run for political office.
- The difference, for Segall, is that in the first case discrimination makes opportunities for welfare more equal, and in the second case it makes them less equal.

But there are problems with using luck egalitarianism to justify affirmative action. Firstly, the support for affirmative action policies that luck egalitarianism affords is highly contingent. Luck egalitarianism supports affirmative action policies only if they contribute to greater equality of welfare overall. But when exactly they are likely to do so is an extremely complex and disputed empirical question. To know when this is the case, we would need to find out a) what the existing distribution of opportunities for welfare is; b) what the exact effect on its beneficiaries’ welfare preferential hiring would have. And c) finally, it needs to be shown that affirmative action would be a more efficient way of equalising opportunities than alternative policies like redistributive taxation.

Secondly, even if luck egalitarianism does support affirmative action, it may also accept forms of discrimination that are intuitively unjust. This is something we discussed in the last lecture. There could conceivably be a society where different two racial groups have roughly equal power, and where each group refuses to hire members of the other group. Here, discrimination doesn’t create inequality of opportunity for welfare, because everyone is equally subjected to discrimination. So, luck egalitarianism thinks this discrimination is fine. But this seems false.

Lastly, even if luck egalitarianism could support affirmative action while ruling out objectionable forms of discrimination, it remains a very controversial ideal: as we have seen before, luck egalitarianism requires eliminating all undeserved natural and social disadvantages. As a result, one might think that it would lead to excessive state interference with individual liberty. So even if luck egalitarianism supports affirmative action, you may not think that luck egalitarianism is itself appealing.

**4. Meritocracy and affirmative action**

Let’s turn to the other theories of equality of opportunity we have studied—meritocracy and fair equality of opportunity (FEO)—and see whether they can support affirmative action. Meritocracy says that desirable jobs or positions should be offered to the best-qualified applicants through competitions that no one is excluded from entering. FEO keeps M Meritocracy, and combines it with a further principle, Fair Background, which asserts that access to qualifications should not be influenced by individuals’ socioeconomic background.

As in the last lecture, I’m going to treat meritocracy and FEO together in the rest of this lecture, by focusing especially on the Meritocracy component of FEO. This isn’t to say that Fair Background is irrelevant to affirmative action. Indeed a defender of affirmative action could try to argue that affirmative action is justified because it helps to nullify the influence of social background on access to qualifications. Here, though, let’s focus on the meritocratic principle that we should choose candidates only according to their qualifications.

Notice first of all that there is something strategically advantageous about seeing if we can justify affirmative action in terms of meritocracy. Meritocracy is the most minimal version of
equality of opportunity that we have seen. It is less demanding in terms of policy changes than luck egalitarianism. And it is presupposed by FEO. Partly for this reason, meritocracy has wide popular appeal. So if we can show that meritocracy supports affirmative action, then this will give us a justification for affirmative action that is persuasive to a broad political audience. As such, even if you don’t think meritocracy is the best principle of equality of opportunity, starting from meritocracy may be an effective way of showing your political opponents that their principles commit them to affirmative action.

Now, at first glance, the problem is that meritocracy and affirmative action seem to be straightforwardly in tension with one another (Miller 1999, ‘Deserving Jobs’, p. 156; Segall 2012 ‘Should the Best Qualified be Appointed?’, p. 36). Meritocracy requires that we select candidates only on the basis of their qualifications—that is, on the basis of their abilities to perform the job. But affirmative action says that the social group to which the candidate belongs—such as their race or gender—can also be used as a basis for hiring. So it seems that affirmative action allows the selection of people who are not necessarily the best qualified. As a result, meritocracy seems to reject affirmative action as discriminatory. If we want to justify affirmative action, how can we respond to this problem?

4.1. Affirmative Action as a Means to Meritocracy

A first possibility is to try to show that although affirmative action violates meritocracy here and now, it helps to realise meritocracy more fully in the future. The main argument for thinking this, which Anita Allen (1990, ‘On Being a Role Model’) has discussed, is sometimes referred to as the ‘role model argument’. According to Allen, because women and people of colour are currently underrepresented in some prestigious jobs and positions, they may fail to aspire for those jobs. In other words, they might have trouble imagining themselves in those jobs, or feel that they wouldn’t belong in those positions. And they may feel this even if they are in fact perfectly qualified for those jobs. The result is that people from these groups who are qualified don’t try to get the job.

Affirmative action counteracts this problem by getting more people from underrepresented groups into prestigious jobs and social positions. These hires serve as role models, who provide evidence to other members of their groups that they too can be successful in those jobs and social positions. This, in turn, encourages those who are qualified or who have the potential to be qualified to aspire for jobs that they otherwise wouldn’t have aspired to. So this helps us better realise meritocracy in the future by bringing about a better pool of qualified applicants to choose from.

Though this argument may seem appealing, it faces important difficulties. Specifically, like the luck egalitarian argument, it depends on contentious empirical claims about the causal consequences of affirmative action policies. If things turn out well, affirmative action will create role models, who contribute to generating a more meritocratic society. But as Anita Allen herself notes, things might turn out less well than that. An important worry is that affirmative action policies might inadvertently send the message that people of colour or women who are appointed are undeserving of their job, and that they are less qualified than others. In fact, Allen worries that the role model argument might encourage that interpretation: this is because it suggests that members of underrepresented groups who are currently hired were hired less because of their qualifications and more because their appointment will encourage qualified candidates to apply in the future.
4.2. Affirmative action as constitutive of meritocracy

Let’s consider a different meritocratic defence of affirmative action, which tries to overcome the problems of the role model argument. Unlike the role model argument, this defence of affirmative action does not say that choosing less qualified participants now will help get better qualified candidates in the future. Instead, it says that affirmative action policies are actually ways of selecting the best qualified candidates here and now. So it denies that there is any tension between affirmative action and meritocracy. Affirmative action is constitutive of meritocracy. Now, this may seem puzzling. As we mentioned earlier, using membership in socially salient groups like race or sex as a criterion selection seems, at first sight, to be inconsistent with selecting according to qualifications. So why think that affirmative action is a way of hiring the best qualified?

The reason, Elizabeth Anderson (2010, *The Imperative of Integration*, ch. 7) has recently argued, is that our judgments about who is qualified tend to be biased or skewed against people of colour and women. There are several grounds she gives for thinking this. The first reason is that the evidence that employers have of applicants’ qualifications may be unfairly skewed. If, as Anderson thinks, we are in a society where many people are bigoted or sexist, our evidence regarding who is most qualified may be inaccurate. For instance: if a female applicant has a reference writer who is sexist, the risk is that her reference writer will underplay her qualifications in her reference.

Secondly, even if our evidence isn’t skewed, Anderson points to the problem of implicit bias (see also Kelly and Roeder 2008, ‘The Ethics of Implicit Bias’). There is growing evidence that prejudicial stereotypes unconsciously influence our judgments about people, even when we explicitly reject those stereotypes. This implicit bias influences our assessments of how qualified people are. For instance, some influential studies asked employers to judge which CVs manifest the best qualifications. When the CVs had names that sounded more ‘white’, employer tended to judge that they were stronger CVs, even though they were in fact identical. (See, e.g., Bertrand and Mullainathan 2004, ‘Are Emily and Greg More Employable Than Lakisha and Jamal?’). The point, then, is that either because our evidence is skewed, or because we process information in a biased way, our judgments regarding who is most qualified tend to underestimate the qualifications of certain socially salient groups.

Now, if we know that our judgments are biased in this way, Anderson thinks, then we should adjust them, by giving extra points to those we are biased against. And according to Anderson, this is precisely what affirmative action policies do: they give preference to groups that employers are involuntarily biased against. “Affirmative action”, she claims, “is an application of Aristotle’s point that to do the right thing in the face of a contrary inclination, we must drag ourselves in the opposite direction, as an archer must aim against the wind to hit the bull’s eye” (2010, p. 148).

Notice two advantages of this justification for affirmative action.

1. Firstly, unlike the role model argument, it does not send the message that those selected through affirmative action are less qualified than others. Instead, affirmative action is seen as a way of choosing those who are actually the best qualified in conditions where our evidence or cognition is biased.

2. Secondly, and relatedly, on this view, affirmative action is not discriminatory. Membership in a racial or gender group is not being used as a selection criterion in
and of itself. Instead, it is being used in order to track the real criterion that we are after: namely, who is most qualified.

One potential limit of this argument is that it too rests on complex empirical evidence. In particular it rests on psychological assumptions about implicit bias. As I’ve said, there is quite a large body of evidence supporting the claim that implicit bias is real. But it has to be mentioned that there are still disputes about exactly how well the test which has been designed to measure implicit bias (the Implicit Association Test) really predicts discriminatory behaviour (for an overview of this controversy with links to research articles, see Bartlett 2017, https://www.chronicle.com/article/Can-We-Really-Measure-Implicit/238807).

Another issue is that even if this justification reconciles meritocracy with affirmative action, you may not think that meritocracy is a very desirable ideal. I’m not going to go into this here, because we’ve already discussed the benefits and problems of meritocracy in the first lecture. But if you’re interested in an attempt at showing that meritocracy is not valuable in the specific context of affirmative action, I strongly recommend reading Iris Marion Young’s chapter ‘Affirmative Action and the Myth of Merit’. There, she offers a highly influential critique of meritocracy where she challenges the claim that there are impartial standards for assessing people’s qualifications.

4.3. Non-meritocratic justifications for affirmative action

Because the meritocratic justifications of affirmative action rest on empirical claims that remain contested, and because the value of meritocracy is itself disputed, it’s worth examining whether we can find non-meritocratic justifications instead. Let’s suppose, then, for the sake of argument, that meritocracy doesn’t support affirmative action. The question then is what values or moral considerations might justify affirmative action in spite of the fact that meritocracy opposes it. Here there are two popular suggestions.

(a) The Compensation View

The first, which has famously been defended by Judith Jarvis Thomson (1973, ‘Preferential Hiring’) and George Sher (1999, ‘Diversity’), is that affirmative action is justified as a compensation for past injustices. The idea is that people who have been wronged in the past are owed compensation from those who have wronged them. And affirmative action just is this compensation. It gives preference in hiring to members of social groups that have been disadvantaged. More concretely, on this view, people of colour may be given preference in hiring as a form of compensation for the fact that they have endured a long history of racial discrimination.

This justification encounters some problems, however.

(1) Firstly, in actual affirmative action policies, all members of a racial or gender group are targeted for preferential hiring. This includes the younger generation, many of whom may never have looked for a job before. So it might seem that actual affirmative action policies benefit some individuals who haven’t themselves been wronged or discriminated against in the past.

(2) Secondly, the burdens of affirmative action don’t necessarily fall on people who have discriminated in the past. Rather, it falls on other candidates for the position, some of
whom may never have engaged in acts of discrimination. Accordingly, one worry you might have is that if affirmative action is there to right a wrong, it is actually holding the wrong people responsible for the past injustices in question.

Now, Thomson herself doesn’t think these criticisms are decisive.
- Against the first problem, she notes that we shouldn’t underestimate the extent to which people of colour and women continue to be discriminated against today. “It is absurd”, she says, “to suppose that the young blacks and women now of an age to apply for jobs have not been wronged” (1973, p. 381). So, on her view, the younger generation too may well have faced discrimination.
- Against the second problem, Thomson suggests that even if white male candidates who are applying for jobs today may not have perpetrated any discrimination, they may nonetheless have benefited from past discrimination. For example, in the US context, the idea is that white Americans historically accumulated wealth partly by excluding black Americans from desirable jobs, and their descendants in turn benefit from this unjust accumulation of wealth.
- Finally, even to the extent that affirmative action policies benefit some people who have not been wronged, and create some costs for innocent candidates, Thomson suggests that this may just be a necessary imperfection. On this line of thought, it’s very difficult to find out who exactly has been discriminated against in the past, and who has discriminated against others. Using racial or gender groups is an imperfect proxy for who has been wronged, but Thomson thinks it remains the best proxy available (1973, p. 382).

One might harbour some reservations about Thomson’s replies. Firstly, you may have doubts about the claim that using race or gender or other socially salient characteristics really is the best epistemic proxy we have. There is huge variation in good and bad fortune within socially salient groups. Therefore, one might worry that membership in a socially salient group is a very imperfect indicator of who has been wronged in the past.

Secondly, and perhaps most importantly, what Thompson doesn’t fully explain is why exactly affirmative action is the right way of compensating people for past wrongs. As Anderson (2010, *The Imperative of Integration*) notes, you might think that financial compensation and public apologies would be a better form of compensation.

**(b) The Diversity View**

The second and final account we’ll consider today of why affirmative action may be justified in spite of meritocracy originates from the value of diversity. According to this line of defence, affirmative action is desirable because it leads to greater diversity in the hiring institutions. We end up with a greater range of social groups being represented in jobs and positions.

The crucial question here is why exactly diversity is valuable. One of the most influential arguments, which can sometimes be found in Iris Marion Young (1990, ‘Affirmative Action and the Myth of Merit’), is that diversity is valuable for epistemic purposes. For Young, “because of their different experiences, cultures, values, and interactive styles, people from different groups often bring unique perspectives to a collective endeavour … [T]he primary purpose of affirmative action is to mitigate the influence of current … blindesses of institutions” (1990, p. 198). In other words, because they encounter different social
constraints, people from different social groups often have different experiences and perspectives. So, having people from a diverse range of social groups in an institution allows that institution to pool or collect very different perspectives. And this in turn allows the institution to make decisions that are better informed.

But this diversity rationale for affirmative action too is not entirely satisfying. Firstly, Anderson (2010, *The Imperative of Integration*, ch. 7) notes that not all institutions need epistemic diversity. Some jobs don’t demand much personal input from workers (e.g., jobs at a call centre). And some jobs produce undifferentiated services (e.g., delivering mail). For jobs like these, it isn’t clear that having a workforce with diverse perspectives would improve how well the job is performed. As a result, the diversity argument wouldn’t support affirmative action policies for jobs like these.

Secondly, George Sher (1999, ‘Diversity’, p. 99) observes that the diversity argument doesn’t explain the special focus on gender and race that we see in actual affirmative action policies. After all, there are many other dimensions of diversity than being from different racial or gender groups: for example, whether one comes from a different country, what political ideology one endorses, whether one comes from urban or rural areas, and so on. So if what we are after is simply diversity, Sher asks, why not also use affirmative action to make sure that we have “suitable numbers of religious fundamentalists … ex-military officers, conservatives, Marxists, Mormons … These groups, too, have characteristic concerns, types of experience, and outlooks on the world. Thus … why not also give *them* preference”? The point, to conclude, is that unless you can explain why the socially salient groups that affirmative action focuses on provide a distinctively important kind of epistemic diversity, the diversity rationale isn’t going to support affirmative action in its existing form.