

[SQUEAKING]

[RUSTLING]

[CLICKING]

**FRANK**

**SCHILBACH:**

All right. Welcome, everyone, to lecture 13 of 14.13. This is my last lecture on social preferences. I hope all of you are doing fine remotely. Hang in there. I know this is a very difficult time. This lecture is a bit of a more uplifting lecture than the previous ones. While the previous ones have looked at social preferences, how we can measure social preferences, and how can we perhaps isolate pure altruism-- do people really want others to be better for their own sake, even if nobody knows about it, either others or whether there's no feedback-- we sort of found out that like it seems like what looks a lot like altruism-- people are nice to each other-- is not pure altruism in the sense that like they really want others to do well.

Instead, they want to either look good in front of others, or they're worried about reciprocity or sort of any other negative feedback from others, or it might be they might want to protect their self-image. They want to look good in front of themselves and they try to figure out ways to deceive themselves that they're nice. And given the opportunity to not be nice, people actually seem to be not as nice as one might think from pure dictator games or ultimatum games that we observe.

This lecture is a bit more uplifting in the sense that it talks about field evidence on social preferences, and in particular you can ask the question about can policies increase pro-sociality in certain ways. We're going to talk about three broad sets of studies. First, we're going to talk about social preferences at the workplace. I'm going to talk about the impact of relative pay on productivity.

We talked about this last time at the end of the lecture a little bit, but I'm going to just restart from scratch to sort of explain that a little bit better. Then we're going to talk about the morale effects of pay inequality. What happens to worker inequality when people are paid unequally, and is that a good idea? There's also a very nice paper in ethnic divisions and production in firms. We're going to talk about this a little bit in recitation.

Number two-- we're going to talk about whether policies can induce pro-sociality. In particular we're going to focus on mixing. This is sort of called the contact hypothesis. When you mix people from different backgrounds, do they become nicer to each other in various ways?

You'll discuss a very nice paper by Gautam Rao when mixing rich and poor children in school. And I talk very briefly about mixing cricket players in India to reduce tensions across castes or make people from different castes perhaps more prosocial to each other. And then I'm going to also briefly mention a study on mixing roommates in college by Corno et al.

Finally, we're going to talk about whether people perhaps underestimate the benefits of prosociality. In particular, I'm going to show you a study on people undervaluing gratitude. So when people write letters of gratitude, do they understand how others might react, and is there perhaps some evidence that people are not prosocial enough because they misunderstand the effects that that might have on others?

OK, so let me start with social preferences at the workplace. So Bandiera et al. is a very nice study that looks at the impact of relative pay on productivity. This is field evidence from the fruit farm in the UK. It's literally field evidence because it's from fruit-picking farms from fruit fields. The study looks at two types of payment schemes. These are piece rates-- workers are paid per unit of output. And the second scheme is relative pay-- workers are paid relative to others.

We might think that relative pay is a good way to incentivize workers because you want to do better than your neighbor, and there's a computation of who does best. And in principle, it could be that that, really, if you only care about yourself, that really gets people to work very hard.

Now, what Bandiera et al. look at is, how does the introduction of relative pay affect workers' effort and output? Now, one thing to notice is that while relative incentives might be good because you might induce people to want to be better than their neighbors or their co-workers, there's also a negative externality of relative pay.

That is the idea that increasing your own pay comes at the costs of others pay. That is to say, if there's relative pay and I work really hard, I'm paid how I do relative to my friends or others that I'm working with. Well, if I'm working really hard and do better than my friends, I'm going to be paid more. But at the same time, for any effort that I put in, my friends are going to be paid less.

And so workers anticipating these impacts on others might reduce their effort if they care about others. If I'm really worried about if there's relative pay and I work really hard, my friends are going to look bad. They're going to be relatively unproductive, and I worry that they might not be paid enough. Then I might reduce my effort. And then everybody might do that, and then the introduction of relative pay or the relative pay schemes might actually not be particularly effective or an effective way of incentivizing and motivating workers. So there's sort of this tension that relative pay in principle could work quite well. But if people really care a lot about others, then it might not work at all. It might actually backfire in some ways.

Notice that it could be either that people care about others for their own sake in the sense that you might care about them because you want them to have high earnings, or you might care about others' earnings, not because you actually care about the outcomes per se, but rather because you're worried about negative repercussions. If these are your friends and you work really hard, well, at nights, your friends might really not be happy, and there might be retribution against you socially or in some other ways if you work too hard, if you caused them their earnings to go down.

OK, so now, what does this paper do? It looks at how does switching to piece rates affect people's productivity? So the study has personnel data from a food farm in the UK. They measure productivity as a function of their compensation scheme. This is a quasi-field experiment.

The timeline is as follows. For the first eight weeks of the 2002 picking season, fruit pickers were compensated on a relative performance scheme. So the per fruit piece rate is decreasing in average productivity. So if everybody works really hard and I don't, if I'm relatively bad compared to others, I get less than others. This is an incentive, as I discussed, to keep the productivity low if you care about others.

And then in the next eight weeks, compensation will switch to a flat piece rate per fruit. So that's essentially just you're paid by however how much you produce, how many kilograms of fruit you collect. And the payment is entirely disjoint or independent of what other people do. That's a classic piece rate payment.

So the externalities are entirely shut down. If I work really hard, I'm going to be paid. There's no effect whatsoever on my fellow workers. The switch was announced on the day the change took place, so it came as a surprise to workers. So we should not be worried about sort of pre-trends or productivity going up over time anyway.

Now, what the paper finds is a dramatic increase in productivity with the introduction of piece rates. That is to say what you see here in the figure is the average worker productivity for two of the fields-- so that's essentially like broad, big fields where people work on, and I'm showing you the data for two of those fields, for which, essentially, on most days, there was actually production workers were picking fruit.

What you see on the left side of the graph is when people were working under relative pay, there you see, essentially, relatively flat productivity. People picked about 5 kilograms per hour. There doesn't seem to be any trend upwards over time. So before the new policy was introduced, essentially, workers were producing or picking their fruit at a relatively constant rate over time, about 5 kilograms per hour. There was no trends, as I said.

Now, you see the vertical line in the middle. This is when piece rates were, in fact, introduced. So there was no relative pay anymore, and now people were paid by piece rate. And that increased productivity by over 50%. So workers became a lot more productive after that. That's true for both of those fields.

It is not the case that the average payment per unit of output per kilogram of fruit paid was actually was going up. So you might say, well, if the piece rate was actually higher under the piece rate payments on the right side of the screen, of the graph, then maybe workers would just be more productive. But in fact, if you look at the piece rate over time, if anything, the piece rates actually went down over time what the company was paying.

So that is to say, the introduction of those piece rates, the productivity effect is not explained by workers being paid more per unit of output. Instead it seems to be really coming from the different incentives people workers have when it comes to how their productivity affects their fellow workers on the field.

Now, you might have two potential explanations here for this evidence. One is social preferences. So you might work less to help others under relative incentives, even less when your friends benefit. You have many friends on a certain field that work with you. Well, then you might be particularly inclined to not work very hard because you care a lot more for them. And this is exactly what they find in this paper. These effects are stronger when there are more friends on the field for a particular worker.

Second, however, it's also like a repeated game. That is to say, there's a low effort equilibrium, where, essentially, if there relative pay, we can just all agree everybody on the field might agree that not working particularly hard is a good idea because we are essentially paid relatively anyway.

So if everybody doubles their effort, nothing is going to happen to a worker's compensation. So you might as well just decide everybody could just decide work half as hard, and you're going to be paid the same. And as long as we can sustain that equilibrium, that's a good idea. Of course, in any given period, any given worker might have the incentive to deviate from this equilibrium because on any given day, if you work really hard, you're going to be paid a lot more for that day.

And so how might we be able to disentangle that? Well, what Bandiera et al. have is they have also some variation in the types of fruits that were collected. What do I mean by that? Well, if you think about what you need to be able to disentangle these, what you might want to have is differences in the observability of effort, right?

So if I can see what my co-worker does, I can punish them very effectively. I can just know whenever Frank is working really hard, my co-workers might want to punish me. And then it's very hard for me to work really hard because I'm getting in trouble at nights after work with my co-workers. Or in the next day, everybody might also work really hard, and I don't want that.

On the other hand, if I can hide how hard I'm working, well then, I might secretly work really hard, and my co-workers might not even find out. So then, if that's the case, in the repeated game equilibrium might just not be enforceable, as in I might just tell my friends, oh, I wasn't really working hard. But in fact, I was. And they have no way of finding out because I was unobservable.

If, instead, I really care about them, if I'm really interested in their well-being overall and their payment, it doesn't matter whether my effort is observable. Regardless of how observable my effort is, I will not work very hard because not working hard makes them better off. So if you just had some differences in observability in effort, how hard people work, we could just disentangle these two explanations.

So Bandiera et al. have this very nice variation. They have two fruits there, fruit type 1, which are strawberries, and they have fruit of type 2, which are raspberries. If you ever have picked strawberries, you would know that strawberries is very easy.

These are essentially very flat and low fruit that when somebody picks strawberries, it's very easy to see how fast somebody works, how much anybody picks and so on because you can just see across the entire field. You can see essentially what all of your coworkers are doing.

In contrast, if you look at raspberries, these are bushes. So you can essentially literally hide behind the bushes and work really hard to pick really hard and fast without others noticing. So essentially, productivity is at least more unobserved than for strawberries.

Now, what Bandiera et al. then find, in fact, is no impact of the piece rate on fruits of type 2, which is raspberries. So that suggests there is no evidence of pure altruism, and the effects, perhaps, could be driven by reciprocity. This is consistent with what we had found before.

When you remember the dictator ultimatum games that we showed you, it looked a lot like people are altruistic. They care about others. But when you look at their actual motives, when you dig a little deeper, we find out that, in fact, this is driven by reciprocity, or by worries about repeated game effects, or worries about essentially just retribution, either on the field that other people will work also really hard if I work really hard right now, or socially after work, my coworkers might be really mean to me. They might even beat me up, or they might not be my friends anymore.

Here is regression tables that show you this. The first column, you see fruit type 2. Again, these are the raspberries. This is when things are unobserved. There's no effective introducing the piece rate. In contrast, if you look at column 2, this is fruit type 1. These are strawberries, and there's large productivity effects of strawberries there.

So these results, again, highlight the importance of setting the incentives carefully. Small details matter, and you really want to be very careful in how you said relative incentives that in principle seemed like a good idea because they provide steep incentives to workers, but in practice might actually not work particularly well.

OK, so the second paper we're going to look at is Breza et al. These are the morale effects of pay inequality. This is a very nice randomized field experiment in rural India with low-skill manufacturing workers. The question the authors ask is do workers care about relative pay? And perhaps why do we see so little variation in pay across settings?

So what we often see in many companies in India, but also in other places, also in certain villages in India, what we see is wages tend to be very compressed. So workers essentially earn the exact same wage, regardless of their underlying productivity, experience, and so on. Good and bad workers tend to earn exactly the same amount of money for per day or per hour and the like.

Economists think that's inefficient because if you want to hire somebody who's really, really productive, you might want to pay them more. If you want to hire somebody who is not that productive, you might want to pay them less. You might actually want to pay them less than the prevailing wage, but you might not be able to do so because the norm is that you have to pay them the prevailing wage. And then you might end up not hiring that person, and that person might end up being unemployed, even though you would be very happy to pay them at a lower wage.

This is a really relevant question. Sorry, the second question the authors ask is what is the notion of the underlying notion of fairness? What is fair and what's not fair? As in if there is wage dispersion across workers within teams or within a certain firm, under some circumstances under which that's OK to do?

Are there some good justification that workers are OK with, or is it just always bad to do so? So can we find some situation under which it's justified for workers or not, or is it just like we cannot pay have pay inequality under any circumstances in those kinds of firms?

This is potentially quite relevant for many features of the labor market. It might explain wage compression. That's what I'm saying, like wages tend to be fairly compressed, even if productivity is fairly dispersed across workers. There tends to be lots of wage rigidity. Wages tend to not move very much. Employers have trouble adjusting wages, or they are reluctant to adjust people's wages much, in particular when it comes to like downward rigidity. We talked about this a little bit already when we talked about reference dependence.

This can help us think about like sorting of workers into firms and inequality. We can think about firm boundaries in a sense and say should they be really large firms or a number of small firms? If you think the relative comparison is across workers within firms, that would say, well, we should probably have a lot of small firms.

And then each of these small firms could have different wages, as opposed to one large firm where essentially everybody has to be paid more or less the same. And we might think that's inefficient because as returns to scale of having a large firm. Think about it is also relevant for some HR policies in terms of how do you pay workers, how do you set wages, and so on.

So as motivation in the study, the authors ask people the following questions. The questions are the following. This is rural India in Rissa, a relatively poor area with lots of small-scale manufacturing. The question is three people from a village get hired to work on a construction site together.

The prevailing wage is 250 rupees. That's the standard wage in the village. Each village tends to have like a prevailing wage, which is how much people are paid usually. The contractor pays them 250 rupees per day. How well will they work together?

And what you see is what respondents tend to say is 80% of respondents say people worked very well together. 20% say as well as usual, and 0% say there will be conflict. So that's seems like paying everybody the same is the socially acceptable thing to do. Everybody seems happy with that.

In contrast, if you ask the following, which is the contractor pays them different wages. If the contractor pays 250 per day, 270 per day, and 290 per day, how well will they work together? And so now, what you see here is when there is differences across pay, 94% of workers say there will be conflict.

Listen, in this version of the question, it doesn't specify whether that's due with underlying productivity differences. So why is it that some workers are paid different than others? However, it's very strong suggestive evidence that workers might be really unhappy, and there will be conflict across workers if they're paid different wages in the same kind of team, or at least on the same construction site.

And then, of course, if you're an employer, you might be quite worried about paying people different wages because of this conflict. Presumably, workers will not work or collaborate as well as they could potentially.

OK, so now the authors set up a very nice field experiment that looks like this. So there's 10 production units of 3 workers in each factory. So think of factories like work sites, factories in quotation marks because it's really not just a factory as you might imagine it. It's really like a small work site, an office where people are working in different areas. There's 10 units of 3 workers each, so it's not a huge office.

Each unit of three workers produces different products. So each unit has one product. For example, unit 1 makes brooms. Unit 2 makes incense sticks. Unit 3 makes leaf plates, and so on. So each unit has one item of production.

Now, why are there different units? It's essentially to separate workers into different lines of work. So what we're going to expect is that workers are going to compare each other within unit, but not so much across units because if I make brooms and you make incense sticks, that's very different work in some ways. But if you make the exact same thing, the natural comparison now for workers to compare yourself with another worker.

All unit members, so everybody in a given unit, makes the exact same product. So everybody in unit 1, who is three workers, makes brooms. Everybody in unit 2 makes incense sticks, and so on. Now the key experimental variation in the experiment is weight dispersion across workers within teams. Finally, there's also-- and I'll get back to that at the end-- production tasks vary in their observability of performance. So that is to say for some tasks it just happens to be that it's much easier to understand how productive your co-worker is than in other tasks. OK, so now what is the variation in relative pay.

So here you see people are paid depending on their worker rank. There's four different regimes of pay in the different columns, heterogeneous, compressed L, compressed M, and compressed H. Workers are sorted. So there's a baseline period where workers who work for a few days, and the few days are used to assess their baseline productivity.

They're ranked into three turnstiles, so there's like the low productive workers, median productivity, and high productivity workers. They're ranked essentially based on how well they did at baseline. OK, so if you're like really bad at the start, you would be a low productivity worker. If you're like an average worker, you would be median productivity. And if you're really like highly productive worker, you would be classified as high productivity workers.

Notice that's always done for all of the different treatment groups. The difference now is across these different regimes how workers are paid. In their heterogeneous pay essentially, there's differences in cross workers in the sense of like the low productivity workers are paid low wage, the median productivity paid workers are paid the median weight, and the high productivity workers are paid the high wage.

As you might expect, that's, sort of, the natural thing to do in terms if wanted more productivity, the most productive workers you're going to pay the most. Notice that these wage differences are modest. Even like the difference between  $W$  high and  $W$  low are like only about like up to 10%. So really these are not huge differences. It's not that the other guy if you are like low productivity worker if you get  $W$  low, it's not the other guy gets like twice as much. It really is like modestly more, but it is more money that they get.

Now then if you look at the other three regimes, compressed L, compressed M, and compressed H, low productivity-- so in compressed L, everybody gets to low wage, in compressed M, everybody gets the minimum wage, and in compressed H, everybody gets the high wage. And now the study now lets the authors compare the different columns for holding exactly constant baseline productivity and wage levels. That is to say, for instance, we can compare the heterogeneous-- the low productivity workers that happen to be randomized into like the heterogeneous treatment, they receive  $W$  low or compare them with like low productivity workers at baseline who happen to be randomized into a compressed L. So I should have said more clearly people are randomized into any of these four groups, and then depending on what your baseline productivity is you get the wages as I show here in the table.

So for example, if you are like a low productivity worker, you might be randomized into the heterogeneous treatment group. You might receive  $W$  low, or you might be randomized into the low-- compressed, low treatment where you also receive  $W$  low. Notice that in both cases, that worker is a low productivity worker. In both cases, the worker receives  $W$  low. However, what's different here now is that his co-worker-- his or her-- in this case, these are all men-- co-workers are receiving either the same, which is in compressed L, or they receive higher wages in the heterogeneous weight treatment, right.

So now we can look at workers who have the exact same baseline productivity on average at least and have the exact same weight, but what's being varied is like how much other workers are earning. We can do this for low productivity workers. We can also do it for median productivity workers or also for high productivity workers. I skip the median productivity worker, but that's exactly the same. For high productivity workers, we can look at the workers who had been randomized into the heterogeneous treatment group or heterogeneous worker group-- sorry, heterogeneous wage group where the worker gets a  $W$ H but everybody else gets like a lower wage, or we can compare that to compressed age where everybody gets  $W$ H in that group.

OK, so now what do you what do the authors find? Let's start with the low productivity workers at baseline. Remember the collinear comparison here is between pay disparity-- this is a heterogeneous group-- pay disparity and compressed L, which is like the group where everybody is paid the same. Now what the authors find-- and you see this in the graph fairly nicely, which is like the productivity-- these are like the red line and the blue line-- the productivity on the left side- this is before the treatment starts for about over 10 days-- the productivity looks pretty much the same.

But then when you go to the right side of the graph, not immediately but after a few days they seem to be a gap or a gap emerges between the red line and the blue line, which is exactly the gap as you would expect if the disparity makes workers less productive, which is to say the workers who receive a low wage but others in that group are receiving higher wages are becoming less productive compared to workers that receive a low wage where others have the exact same productivity.

And you see this also in the regression tables, you see about a 22% reduction in mean output and a 9% reduction in earnings, which are pretty large effects. So these are like large effects compared to like other interventions that people have tried. Interesting, you see a little bit like while the treatment effect when you look at the graph initially looks not particularly large, if anything it might not even be there often like a few days, that treatment effect increases over time. It becomes larger so over time people become less and less productive compared to the compressed wage treatment.

Interestingly, we find-- the authors find similar effects for high ranked, high productivity workers. So notice that these are not workers who either who are high productivity workers to receive the same wage in the pay disparity treatment and the compressed H treatment that have the same wage. In one case, they are paid exactly the same. In the other case, they're paid more compared to their coworkers. And what happens what seems to be the case is that the compressed high paid workers are, in fact, more productive compared to others.

You might have expected in some ways like if one guy gets paid high payment but then others get paid less, that makes the high pay worker more productive, because maybe he feels good about himself that he's a high paid workers. And maybe he feels he can prove himself or the like, but instead what seems to be the case the group that works in his team that does not work well together or they just becomes uncomfortable to work with somebody else who is like really mad at you for earning more than they do. And so now the high pay workers in the compressed treatment, in fact, earn more or produce more than the high paid workers in the pay disparity treatment.

So that is to say the pay disparity, pay inequality does not only reduce worker's productivity for the low pay workers, so not only the people-- the workers who are earning less compared to their co-workers, and that's in some sense to be expected. You might just be mad at everybody else or other people in your group are earning more. Then you might be unhappy and mad about that and just then produce less. Instead it seems to be the case even the high productivity workers, the workers who earn more than others in their group are producing less compared to the control group where everybody earns the same.

So what did we learn from that? Well, pay disparity lowers worker performance for all team members, and so the interpretation of that is that pay disparity undermines workers ability to cooperate in their own self-interest. The paper has some additional evidence with some cooperative tasks where essentially workers are worse at cooperating with each other. It seems to really be that workers are not happy to work in the same place with somebody else who earns less.

Importantly, the perceived justification is essential in mediating these effects. That is to say I told you previously that for some task it was easier to see who was more productive than others, especially at baseline just because in the task the difference is across highly and less highly productive. Or high and low productivity workers was just wider, and it's just easy to see that Frank is really bad at this and somebody else is really good. And for some task, it's really easy to see that, for others not.

So if it's really easy-- if there's a perceived justification for workers, if workers, sort of, thought, OK, it's really easy to see who is more productive than others, the effects are much weaker than when there is no receive justification. That is to say workers themselves are saying, like, if I'm saying, look, there's this other worker who is really productive and if I'm sort of saying, hey, look, this is really obvious that they're more productive than I am, it's only fair that they're earning more, then the affects are much damage-- then there's much less than an effect that's accepted, and there's much less of a productivity effect, if any, if there's pay inequality.

However, if it seems to be the case that they're, kind of, like producing the same thing, if we like similarly producing the same thing and we're equally good at it or approximately equally good and one person is paid more than I am for no good reason, so it appears that workers are really not happy about that, then work performance, in particular output falls.

So what are the implications? Well, this evidence suggests that weight compression may be more likely in some settings than in others, right. That's to say like if it's easy to justify, if the production function or like the production process is such that all aspects of the production and the output or the performance are easy, measurable, and observable for workers, then workers might say it's OK if some workers are earning more than others. However, if that's not the case, if there's a bunch of stuff in the work process that's really not observed-- maybe it's verbal or maybe just really hard to justify or hard to sort of assess in some tangible way-- then workers might be really unhappy and that might lead to weight compression, because the employer might anticipate that.

So overall, this, sort of, says that like relative comparisons matter quite a bit. Fairness matters quite a bit, and what's really key for an employer and anybody who sets incentives is to take into account these fairness considerations, and one wants to really understand what is fair and what is not. And upsetting workers by violating these fairness considerations or norms might be really, really costly for an employer. So the good news here is that if one understands that well, one can really sort of produce increased productivity quite a bit, and in some cases weight inequalities actually fine. One just has to be careful in figuring out when exactly is that the case, OK.

As I said, there's a very nice paper by Jonas Hjort on ethnic divisions and production in firms. You're going to talk about that in recitation.

OK, the second part of this lecture we'll talk about policies to increase pro-sociality. And so the first of these papers is very nice recent paper by Gautam Rao that looks at the question are social preferences malleable. Like, what are the origins of social preferences? Why is it that some people appear nicer than others? And then once we understand that, perhaps, we can also understand what policies, if anything, can affect social preferences.

The main question that this paper asks is how does being mixed with poor students in school affect the social preferences of rich students. Now in a lot of the, kind of, policies where poor students are mixed with rich students, you might ask the question, how does it affect the poor student. If a poor student for whatever reason would not be able to afford like a rich school or rich student's school and you, sort of, allow them by giving them scholarships or other policies, often the question that people ask is like what are the benefits of doing that. Is the poor student doing better in school, or do they have different types of friends, different networks, and does it lead to better jobs and so on and so forth?

This question, this paper ask a different question. It asks the question about what is the effect of rich students if they randomly or quasi randomly are exposed to being in class with additional or some more poor students. The paper exploits a policy change that introduced an admissions quota of 20% for poor students in primary schools in Delhi. In Delhi, these are rich primary schools.

The paper looks at two sources of variation. There's variation across classrooms that allows the author to look at the overall effects. So you can look at, like, within schools there's going to be treated and controlled cohorts, because the policy was introduced at some point. So for some students, they enjoyed the benefits of being in class with poor students or the costs. We'll see about that.

And for other schools, they're already too far advanced. They were like essentially before the-- they went to school before the policy was enacted, and therefore they were not exposed to having poor students in school. For these treated schools there is essentially variation within school across cohorts that we can look at.

Second within cohorts. We can look at treated and control schools. So some schools were getting additional poor kids and others did not. I'll tell you about this in a second. And then there's variation within classrooms to allow us to look at the role of personal interactions. There are, sort of, idiosyncratic assignments to study groups, so some students happen to be in study groups with a poor kid, and others were not.

Now what does this study measure. The study measures three broad set of outcomes. It measures pro-social behavior and generosity. This is very much what we have already looked at.

In particular, it looks at dictator games and volunteering for charities at school. The paper has a very nice mix of laboratory outcomes, sort of, measured in the field, which is, sort of, the dictator game that you're all familiar with. And field outcomes, were just, sort of, trying to collect real world outcomes in the sense of like things that are, perhaps, somewhat less contrived from the perspective of kids in school.

And what's very nice about the study is that it seems to be that the results are very much aligned between, sort of, these live outcomes, the dictator game type outcomes, and the field outcomes, which is volunteering for charity at school in this case. Second, the author looks at discrimination in social interactions. In particular, he does it all sports contests and looks at teammates selection among these students and then willingness to attend play dates with poor students.

Finally, the author also looks at academic outcomes, in particular test scores and disciplinary infractions. Now why might one want to look at academic outcomes, and why we're interested in social preferences, there's several reasons for that. In particular, an important reason here is that if you're against this type of policy, I might, sort of, show you that, oh, adding poor children might affect the social preference of rich children and so on in various ways. Maybe discretization goes down, but perhaps it's the case that all comes at the cost of academic performance, right.

If you're trying to, sort of, implement this policy and persuade policymakers, maybe teachers or parents of like, OK, let's have more poor students in your school. Well, parents might, sort of, be OK fine. There's going to be like some change in social preferences, some change in discrimination, but really what we care about is test scores or discipline at school, and these poor kids might not be-- might be worse in terms of test scores, and they might have negative peer effects in terms of test scores but also in terms of disciplinary infractions. And, sort of, as a policy question, then it's really important to understand if there are some benefits in terms of poor sociology or reduced discrimination, do these benefits come at the cost of reduced academic performance?

So now what is, in fact, the policy innovation in Delhi in 2007. There was a 20% admissions quota in private schools introduced for poor students in some of these private schools. There's a household income cutoff of \$2,000 per year. So these are not, sort of, the poorest of the pole of the students who are qualified. Schools which received subsidized land from the government were essentially included in this policy change. That's over 90% of elite private schools. So think of these elite private schools.

These are really, sort of, like very rich kids or parents of these kids. There's no fees for the poor kids, because they would not be able to afford these fees anyway. There was importantly also no tracking. So it wasn't that high ability or high performance kids would do very well where they like tracked into the high tracks and the good tracks verses lower tracks. Instead everybody was mixed together, poor and rich, high performing, low performing students.

And the poor kids were selected using lotteries, which in principle also allows the author or others to look at the effective on those poor kids, the students who are selected versus not. To be clear this is not the subject of this paper. Just to give you a sense of the magnitude of like the mixing, you see is that the average beneficiary was a, sort of, at the 25th percentile of the income distribution in Delhi, and wealthy students are very much on the right tail-- sorry on the right tail of the distribution, something like the 90th percentile, 95th, which is about-- the US equivalent of that would be about \$200,000 per year.

The average beneficiary had like the US equivalent income of like \$23,000. Again, these are not the poorest of the pole but relatively poor as about like an order of magnitude poorer than the average person in the actual school. Now this policy then induced large variation across classrooms. If you look at poor students in the fraction of the numbers of poor students in by grade in 2011, I told you the policy innovation was in 2007.

So if you look at 2011-- that's four years later-- anybody who is in grade four in 2011 has essentially no poor student in that class in those rich schools. So for students who were in 2011 in grade four, five, or six. That policy came too late. They did not have any poor classmates.

In contrast, if you look at the lower grades, grades 3, 2, 1, and 0, which essentially preschool or minus 1, which is preschool grades, there's lots of additional poor students now in those rich schools. Now again, there's variation within schools and across schools. The variation within schools is across classrooms. That's, kind of, what I'm showing you here.

But it's also a variation across schools. So just to be clear, the variation within schools across classrooms is like in the treated schools there's some classes like the fourth grade have like essentially zero poor kids in their classes, and the third grade, in contrast, have lots of poor kids in the classroom. They can do that comparison across classrooms within school.

In addition, has also variation across schools. There is, in particular, there's treatment schools, which comply in 2007 as they were supposed to, but then there's also delayed treatment schools, which complied in 2008. So the jump shifted by one cohort. So the cohort, the jump that I showed you here, this is for the treatment schools that complied or this for everybody that complied in 2007.

The delayed one complied a year later, so you'll see them essentially, somewhat later increasing-- or one year later exactly increasing the fraction of poor kids in their classroom. And then there is the control schools that were not subject to the policy, at least until 2013. So they had received-- they had received land from federal government or private foundations. So essentially there was no treatment. That is all students are rich in these types of classes.

So that is-- the land receiving is essentially just-- that was like a fairly arbitrary rule in some ways. So that, sort of, leads to at least quasi random variation. And some schools received the treatment, and others did not.

Now let me show you some outcomes. So one outcome is our friend the dictator game which you have seen quite a bit and played already as well is students are allowed with 10 rupees, which is not very much-- that's about like \$0.15-- and they choose to share-- notice that these are pretty young kids, so 10 rupees is quite a bit of money for them. They choose to share an amount as before between 0 and 10. And Gautam sets this up such that students could-- it was not just about the money, but they also can exchange the money for candy later, add 1 rupee per piece. So that was a pretty good deal.

Then the dictator games for the order was randomized. The kids would play two games. Again, so all rich kids would do that. They play first game one where the recipient is a poor student in a school for poor children. And in game 2, it's a rich student in a private control school. Notice it's a control school to be clear, so it's not a kid that they might know.

So the kids are chosen such that the kids were given the names and the photographs of the school shown to the subject, to the kids. So you could sort of see-- and this is verified in the briefing by the author-- is that the children understand very well who is a rich kid and who's a poor kid, but these are not kids that they would actually know. And something that's a less interesting question if you have gone to school with somebody for a long time, you might have seen them around in your school yard, you might be nicer to them or less nice to them

That's not what the study is looking at. The study is looking at other kids in other schools that look rich or look poor both in terms of the way they're dressed and so on, but in particular the way their school looks like. So essentially these are-- now we're looking at social preferences towards other people, people that these kids don't know.

OK, so now what do we see. So this is a generosity to the poor. So like this is essentially the results from game number one. Again, remember these are all rich kids who are giving now to poor kids, and the question now is being exposed to or does being exposed to poor children for several years, does that affect giving in the form of dictator games towards poor children to start with. And we're going to talk about rich students after that.

So here you see the generosity to the poor in control schools. You see this is the percent to the poor recipient. It's remarkably similar to what we've seen previously, which is about 20% to 30%. About 25% is given to the poor kid in the control groups in the control schools. This fraction seems to trend up a little bit over time, but it's also quite constant over time across grades. You see here by grades two, three, four, and five on the x-axis. On the y-axis, it's the percent given to the poor recipients.

Now adding the treatment schools, now we find that for two grades-- for grade two and grade three, remember the graph that I showed you here. You see that in grades three, two, one, and so on, these are the kids who are exposed to the poor children in class in contrast to grades four and five that have not been exposed to poor kids in their class.

So what this graph shows is that exactly the grades three and two that have been exposed-- in the treatment groups that have been exposed to poor classmates, these are exactly the students, the grades where giving is higher towards the poor compared to the control groups but also compared to the older cohorts four and five, the grade four and five. When you look at grades four and five, they see essentially no difference between treatment and control schools. So they seem to be very similar.

In contrast, if you look at grades three and through where the treatment kids are exposed to the treatment by having the poor kids in their classrooms, these are exactly the years where giving goes up and goes up by quite a bit. While the average was, as I said before, about 25%, that goes up to about 35%. So that's a pretty large relative increase.

Now in addition, as I told you, there are also delayed treatment groups, treatment schools. For those schools, if you were in grade three, it was also too late for you, because the treatment was introduced only a year later. So if you were in grade three, you did not have any poor kids in your school. Now if you look at the yellow line here, the yellow line looks very much like the green line. That is for the delay treatment schools, there's no impact either.

They look very much like the control group of schools, as you would expect. However, in the delayed treatment schools in grade number two, that's the grade where the students are exposed-- the rich students are exposed to the poor kids using the essentially similar effects to the treatment schools, as you expect.

So that's fairly compelling evidence that really the differences that you see across these schools are driven by the treatment that was-- the timing of the treatment as opposed to other potential effects due to like selection or maybe these schools are different and so on. We see essentially no differences in grades four and five for any of the treatment and control schools. You also see no difference in grades three for the delayed treatment schools compared to the control group.

In contrast, we see clear differences in grades two for the delay in the treatment school as compared to the control schools. We also see a clear difference for the treatment schools compared to both the control schools and the delayed treatment schools as well. So that's fairly compelling evidence that game or play, dictator games changes when students are exposed for quite a long time to poor kids in their classroom.

OK so now in addition, we will have variation within classroom. So what I show you here, this is all variation across classrooms. That is to say, if you have a poor kid in your classroom, you will behave differently, or that will change your attitude toward your giving and dictator games towards the poor or other poor children.

Now, in addition, the author also has evidence of variation within classroom. Now, he lets you think about for a second, why is it important to also have variation within classroom? Isn't the evidence that we have here already compelling enough? Why do we need more? And you can think about it for a second.

The reason for that is that the diff-in-diff, the difference in differences comparing essentially across classrooms, identifies the overall effective having poor classmates in class. That might be the result of personal interactions. But it might also be the result of the teacher.

The curriculum might change. Your parents might change in some ways in just telling you, look. There's these poor kids in your class. The teachers might tell everybody, oh, you need to be nice to poor kids. The curriculum might change. It might sort of include things that tell you to be nice to poor kids and so on. You might be sort of exposed to just different types of material.

So really, given the evidence that we have here, when you just compare across classrooms, you cannot disentangle whether this is driven from being personally exposed in terms of interacting with poor children personally by just talking to them in certain ways, or being exposed to them, or learning about them and so on. So you cannot disentangle personal interactions from teacher or curriculum changes that are essentially at the classroom level.

So in addition, then, Gautam also has a within-classroom strategy that exploits the assignments of study groups. What is that doing is essentially, it isolates the role of direct personal interactions and is not subject to sorting concerns. Another concern might be that you might sort of say, oh, well, in this school, they're having poor kids that are allowed into this private school. I'm going to send my kids to somebody else.

So that's also an issue perhaps potentially in the evidence that I showed you before across classrooms. The within-classroom strategy is now looking at everybody who is in that classroom and looking at some kids were just randomly assigned to be in study groups with poor kids, and others were not. And that sort of allowed us to disentangle or allows us to isolate the effect of personal interactions.

So one hour a day, kids were working in small groups of 2 to 4 students. Remember, this was for a very long period of time. So it's one hour per day, but it was also for quite a bit of time. It's not just doing that once or twice. It's more an extended period of time.

And crucially, some schools use the alphabetic order of first names to assign study groups. So that's essentially exogenous variation in personal interactions. Other schools frequently shuffled the groups and only-- which essentially sort of-- everybody was equally exposed to those kinds of-- to different students.

So you are sometimes in study groups with rich students only and sometimes with poor students only. In the first group where the alphabetic order was used, you were either sort of in the study group where you happened to be next to somebody who's a poor kid because they happened to be next to you in the alphabet, or you're not, and that stayed the same.

Now, what does this let us look at? Essentially, the alphabetical order predicts study partners. So now you might say people with different names might also be different. So it could just be that like my name-- because my name is different, I'm more likely to be next in the alphabet to a poor student. For example, if I'm a rich kid and there are some names that are more similar to poor kids' names, it could be that my parents are particularly tolerant. They also gave me a name that sounds like a poor person. And therefore, I'm more likely to be in a study group with the poor kids.

And that's not the impact of being in the study group or rather sort of selection. It could be that names are different or people with different names are different. So that's why having this control group is really nice. So what we can look at now is to say, you can look at kids that have names adjacent to rich students and kids that have names adjacent to poor students.

So if you look at the left side, this is essentially looking at students that have their names adjacent to rich students only or students that have names adjacent to also some poor students. When you look at that, where the alphabetic order was not used, the outcome here is what's your share of having poor study partners? And so there, essentially, there's no difference, as it should be, because the alphabetic order was not used.

There was a shuffling around the kids all the time, so it didn't really matter what your alphabet was. It didn't really matter. But crucially, we can still look at the alphabet. We can look at kids that have names where, Jason, in the alphabet, there's only rich kids, rich students. Or they can look at kids that have names where, adjacent to the alphabets, there are poor students.

And, here you see on the left side of the graph, you see essentially no difference across these groups. So they were equally likely to have a poor kid in their study group. In contrast, on the right side of the graph, you see the alphabetic order was used to assign study groups. So if you had a name adjacent to a poor student in the alphabet, you're very, very likely to have a poor kid assigned to you or be in your study group.

Not always because sometimes, these are groups of three. So sometimes, you were lucky or unlucky depending on how you view it. The group was just above you or below you in the alphabet. But most of the time, if you have a name adjacent to a poor student, that student or any student would be in your study group.

If you didn't have a name adjacent to you or if you only had rich students adjacent to you, there's still a chance that you have a poor student in your study group because it could be just like not the person next to you but another person further down is, in fact, a poor student. You end up in that group with that student. Remember, these are groups of 2 to 4 students. So it could just be that you're in a group of 3 to 4 students, and not the person adjacent to you but the person after that ends up in your study group. So you're still likely to have-- you have still like about like a 40% chance of having a poor kid in your study group.

But there's a huge difference between the two types of groups. The name adjacent to the study group has a huge fraction, about 90%, versus 40%. So crucially, now, we can compare for both types of schools, for the schools on the right-hand side where, essentially, the names are very predictive of having a poor kid in your study group, versus on the left-hand side where the names are not predicative at all whether you have a poor kid in your study group. And so we can-- that way, we can keep selection the same, the types of name are the same. We are isolating the impact of having a poor kid in your study group.

Now, what does the author find? He finds that-- this is what we showed you before, which is in the control group, we had about, in the dictator game, these kids gave about 27%. Having a poor kid in your classroom. That's the evidence that I already showed you. The treatment effect is about 12 percentage points. That's fairly large. That's almost like a 50% increase. A little bit less than that. So that's a huge increase.

And then in addition, when you look at kids that have a poor study partner, notice that these are things not necessarily additive, but if you have a poor study partner versus not, there's an effective about 7.5 percentage points. That's a pretty large effect. So both of these things seem to matter. Personal interactions seem to be quite important.

But in addition, there seem to be also some additional effects perhaps coming from the classroom level. You might-- so there's two types of effects that are possible here. Some types of effects could be from the teacher, parents, curriculum, et cetera. Might be different. Or other things could be just like, even if you don't have a poor kid in your study group, you might still play with them or be friends with them or just seeing them around in the classroom might affect you. But crucially, the personal interactions seem to be very important as well.

Now, one question you might ask is, well, so far, I showed you generosity towards poor children. So that was game number 1. We looked at what happens in the dictator games when a rich kid plays with a poor kid and the rich kids become nicer towards the poor kids.

Now, how about generosity towards other wealthy children? Well, it turns out that that increases as well. So it's a smaller effect size, but it's still substantial, and, in fact, statistically significant. So why might that be? You might sort of ask, well, what's going on here?

So here's the evidence. You sort of see the change in giving to rich participants. This is sort of like a distribution. You see this is the share giving to the rich participant. It seems to be what happens-- these are essentially the fraction of students in the different treatment versus control who give 0%, 10%, 20%, 30%, 40%, 50%.

And what you see-- there is about a 10 percentage points decrease in the fraction who give 0% and about an 8-something percentage points increase in the fraction who give 50%. So it seems to be that there's quite a few students who move from 0%. About 10% of them move from 0% to 50%. And 50-50 is kind of like the equal allocation. So rich kids also becoming nicer towards rich kids in these dictator games.

So what's going on here? There's different potential explanations. Perhaps the most plausible explanation-- the study has some evidence on that. At the end of the day, it's hard to sort of entirely nail this or rule out all other potential explanation, but it's quite plausible that what's happening here is that inequity, students, when you're exposed to poor children, you are essentially a bit more averse to inequality or inequity across people, even in things, in fairly trivial things such as dictator games.

Now, I told you before, that's a little bit funny because dictator games obviously are very narrowly framing people and looking at essentially very narrow outcomes. If we have, like, a 50/50 outcome in the dictator game, that doesn't mean that our life is the same. You might be still much richer than I am. So having sort of equal outcomes in the dictator, the 50-50 allocation in the dictator game, might be narrowly framed, in that particular game, fair. But of course, it's not fair in the grand scheme of things.

But what seems to be the case is that-- and there's some other evidence in the paper that you can read if you would like. It seems to be the case that rich children become more averse to unequal outcomes in the world in general because they essentially see these poor kids who are very smart and are disadvantaged in terms of various ways from having lower wealth of their parents.

The rich kids become sort of more adverse against that, and that translates even into dictator games with other rich kids in a very sort of minor thing in the world and in these dictator games, again, even though these dictator games actually don't change very much. But they really seem to be averse to inequality of these outcomes and move them from the 10-0 allocation to 50-50, even with these rich kids.

OK. So that was evidence on social preferences as measured by dictator games. But people also had some evidence on discrimination. In particular, a small field experiment on team selection. So what does this study do? So its subjects are, again, students from two elite private schools. So now, it's like two of these selected schools are selected.

One is a treatment school. One is a control group. And in addition, Gautam invited athletic poor students from a public school. Importantly, he said, athletic students-- these are students who are better at sports than the rich kids. You might think the poorer kids are better at sports than the rich kids anyway, which is probably true. But now, these are particularly athletic students on purpose who are invited to attend a sports event as well.

Now, students in this experiment, in this game, must choose teammates to run a relay race. Now, when you're a rich kid who is thinking about who should be in my team, you can either choose a rich kid who is kind of like similar to you in social ways, or you can choose a poor kid who you don't know very much. You might actually not like the poor kids. But the poor kid is a lot better in the running, so it might be much better for you. You might be much more likely to win in the game because now your partner in the relay race is much faster. So it's a very nice trade-off between ability-- choosing the fast runner versus social similarity.

And now, what Gautam then is doing is like, if you sort of choose the rich kid, that's then a measure of discrimination because essentially, you're choosing a worse runner in favor of or because you want more social similarity. You don't want to hang out with a poor kid. Instead, you choose the rich kid, which reduces your chances in the race but increases your time spent with the rich kids compared to poor kids.

Let me tell you a little bit more detail of the experiment. Stage 1 is randomization. So the people were randomized to sessions with varying stakes. There's 50 rupees, 20 rupees, 500 rupees per student for the winning team. This is a lot of money compared to students usual pocket money, right? So they would get something like \$10 or 500 rupees is like one month's of pocket money. So that's really high stakes for these kids.

The price is varied, essentially, because it lets us price out. It gives us a price of discrimination. Lets us understand how much are students willing to give up in order to not have to-- or in order to be able to socialize with the rich kids compared to the poor kid. There was a brief mixing to start with to judge socioeconomic status. So the kids were allowed to mingle a little bit. That would allow them to fairly easily understand who was a poor kid and who's a rich kid.

OK. Stage number 2 was ability revelation and team selection. So you could essentially observe a two-person race. Usually, it's one poor and one rich students. Neither is from the old school. So these are not students that you would know from school anyway. But the uniforms make the school identifiable. You kind of know who is the rich kid and who's the poor kid.

Now, then you can pick which of these two runners you want to have as your partners. So that is to say you see them run one by one. It's very easy to see who's faster and who's not. And so again, discrimination here is interpreted as picking the slow runner. So if you pick the slower runner, then it must be because you like something some other characteristics about that person more. The obvious characteristic here is that it's most likely going to be that kid is rich.

OK. Then stages 3 and 4 are the choice implementation relay race. So students are randomly picked to have their choices implemented. So some of those choices-- this is, again, the strategy method. Some of these choices were actually randomly implemented. So there's plausible deniability for the students in the sense of you could just happen to be randomized or you happen to pick some students versus another.

It could just be by chance that you are with one student versus another. So you could-- your freedom to choose essentially provide you cover. Sometimes, as we discussed before, the computer is choosing, so you always have an excuse. So it's intended to reveal student's true preferences as opposed to perhaps what they think that their friends want them to choose or the other runners. Then the relay races were actually held and prizes were distributed as promised.

Number 4, crucially, there was a social interaction. So if you picked your teammate, you actually had to spend two hours of playing with a teammate. Board games, sports, playgrounds, and so on. Importantly, this was preannounce. So now, again, as I said before, when picking your partner, you have the choice between either picking the fast, poor kid, which will really increase your probability of winning, or the rich kid, who is kind of slow and will reduce your probability of winning.

But if you pick that poor kid or the rich kid, you have to actually spend two hours with that teammate playing board games, sports, playground, and so on, and you might not want to do that with a poor kid. OK. So now, first, what is the demand for discrimination? You can look at this graph. It shows very nicely for the different prices. As I said, 500 rupees for winning the race, 200 rupees, or 500 rupees.

If you look at 500 rupees, this is again one months of pocket money. There's no difference between treated and untreated classrooms. That is to say-- so there's 10% of people are, as you want, discriminating the poor. So this is less than 10% is about 7%, 8% of students pick the rich kids, even in the really, really high stakes race. That is to say, when the stakes are 500 rupees, there are, like, about 6%, 7%, 8% of students who still pick the rich kids, and they sort of take into account the chance that that might lose them the race. But they don't have to spend two hours with a poor kid then socializing.

Now, when you look at the lower prices, the fraction, as you expect-- this is the red line. Sorry, this is the green line, the upper line. The fraction who choose the rich kids, the fraction who were discriminating against the poor, increases as you expect. So now, it becomes cheaper. The race is only 200 rupees or 50 rupees. You might be more inclined to pick the rich kid because you know the value of socializing stays the same, but the costs of picking the rich kid, the cost of losing the race, potentially at least is reduced. If you look at 50 rupees, the price of 50 rupees for the game, for the relay race, that's about almost 40% of students now pick the rich kid, even though that might lose them the race.

And now, crucially, in red, the dashed line below, you can see the treated classrooms. And what he finds is that for 500 rupees, when the stakes are really, really high, there's no effect. Essentially, it doesn't really matter whether you are in a treated or a not treated classroom, in part perhaps because there's not much room for going lower than that. So essentially, there's no effect on there because you know the stakes are really, really high anyway. There's not much discrimination. And having had a treated-- having had a poor kid in your classroom doesn't really change that.

But then, very clearly, for 250 rupees and 50 rupees, the lower stakes, there's a clear difference between the green and the red lines. That is to say, there's a lot less discrimination towards the poor kids. Poor kids are a lot more likely to be chosen when a student has-- in treated classrooms when somebody had a poor kid, another poor kid in that classroom for several years. So that's to say there's-- being exposed to these poor children reduces discrimination among the poor among the rich students subsequently.

This is sort of the same graph that I showed you before. Again, that's sort of consistent with what we have before. It gets a little messier than we had seen previously. But similarly, even within classrooms, if you look at grades 2 and 3 versus 4 and 5, the effects seem to be concentrated more pronounced in grades 2 and 3.

Now, finally, as I said, we look at test scores and discipline. So arguably, there are some positive effects on social preferences. And as I said before, the policy question now is, does that come at the cost of academic achievement in some ways? So is it that the rich kid now, by being exposed to the poor kids, may be somewhat nicer and more friendly and less discriminating against the poor? That's all and good. But is it the case now that test scores go down?

So Gautam finds no effect on aggregate test score index or zero effect in Hindi and math. There's a little bit of a reduction in English scores of 0.9 standard deviations. That's marginally significant. That's suggestive but not perhaps particularly large. So these effects are not particularly large. And in particular, in aggregate, so if you agree to everything together, there don't seem to be any significant effects. So perhaps the English scores are suggestive but sort of not particularly large.

There's some mild effect on discipline. Interestingly, there's an increase in swearing. You think that's good or bad. You can think about that. But there seems to be a little bit of an effect in terms of language uses. There's no effect on violent and disruptive behavior, which you might think is a lot more damaging, potentially.

OK. So just summarizing, what does the paper find? Well, having poor classmates makes wealthy students more prosocial and generous. They're more likely to volunteer for charities. I didn't show you that evidence, but that's another piece of evidence that he finds. They're more likely to give in money in dictator games to give them more higher fractions of their shares in dictator games.

They also choose more equitable outcomes in sort of disinterested third party games where essentially, you choose between two other students and their allocation. So it's not just that they're more likely to be willing to give up money that others get, but also, they're more likely to choose equal allocations in third party games and disinterested games.

Again, I didn't show you that evidence. But it seems to be that what's increasing here is sort of like inequality aversion in these sort of disinterested dictator or the types of games, where essentially, these students, by being exposed to poor kids, are now more averse against unequal outcomes in these types of games.

Second, there is less discrimination and more higher willingness to socialize with the poor. They're more likely to choose poor teammates in sports contests. They're also more willing to attend playdates with poor children. Again, I didn't show you that evidence. And then there's some small, negative effect on academic outcomes that I think are mostly negligible.

Now, there's other work on the contact hypothesis that I showed you. So the contact-- what is the contact hypothesis? The contact hypothesis goes back to at least Allport in 1954, which is the idea that interpersonal contacts reduces prejudice until certain conditions. And not just prejudice, but also changes attitudes and social preferences potentially.

So Matt Lowe, who was a PhD student here at MIT, has a very nice paper that looks at cricket tournaments and asks the question whether cricket leagues in India can increase cross-class interaction in pro-sociology. So what he does is he randomizes, essentially, cricket leagues and teams in cricket leagues where people across different castes are now more or less likely to play with each other, both in terms of-- he varies or there's variation in the study within teams.

So are you more or less likely to have a-- or some people are more or less likely to have a lower or higher caste. So a person from a different cast in their team. And he has variation in what he calls adversarial contact, which is they're more or less likely to be exposed to players from other teams in higher or different castes from themselves.

So if you're on the team and have a person from a different caste on your team, you're on the same team. You share the same objective, and you want to win together. So now, having somebody from a different caste or just somebody who's different in various ways in your team might make you like them better. You might be sort of more positive about them. You might be sort of more likely to talk to them.

You might learn about them. You might sort of see some sides in them that you hadn't seen before. So you might be more like they sort of empathize and look at sort of nice characteristic of them and sort of update them positively about people from other tasks, and that changes your attitudes towards them in general.

If, however, you play against somebody from different castes, you might, actually-- that might, actually, if anything, backfire, because you really don't like your opponents. You might see them very negatively. You might be aggressive towards them. You might be unfriendly towards them. You might sort of not like that they win against you or the like. So I'll just add, these adversarial interactions might actually backfire in the sense that they might not foster integration but actually sort of make things worse.

So Matt runs this experiment and finds evidence of increased cross-caste interactions. So people are more likely to be friends, more likely to hang out with others. They also are more generous in dictator and other types of games.

They also are more likely to engage in trade or in economic exchange. So what Matt does is he sort of randomizes gloves. It's like left gloves and right gloves. And he does the same for flip-flops. Left flip-flops and right flip-flops. And people are more likely to trade with somebody from another caste if they ever been on the same team with people from other castes.

If they have a higher fraction of people on their team of people from other castes, they're more likely to engage in all these behaviors. They're more likely to have cross-class interaction. More likely to be prosocial. More likely to engage in economic exchange with people from other castes.

So that's all true for collaborative contact. That is to say, that's contact with people on the same team. In contrast, for adversarial interactions, when people are on the opposite team, having more people from other castes or being exposed to more people from other castes does not have these positive effects and, for some of these outcomes, have even negative effects. So if anything, that sort of backfires. It doesn't-- just-- a mere exposure to others, if you are sort of in an adversarial contact situation, does not really foster prosociology or any of these types of integration. If anything, it backfires.

Now, why is that important? If you think about like, for example, attitudes towards immigrants, it really matters hugely what types of contacts people are exposed to. If people have worked together in the same team, if they have, perhaps, team pay, if they were to work towards the same goal, really, it seems this evidence suggests can foster prosociology and so on and so forth that leads to integration, reduce discrimination, and so on and so forth.

So it's sort of the incentives are aligned or if you could sort of set up incentives that are aligned, people might become nicer to each other and interactions might be fostered. In contrast, if contact is adversarial, if you're worried that immigrants are taking away your jobs, being exposed to immigrants might just do the opposite. So you might see a lot of immigrants. But in a way, if you feel like you're in competition with them, if they're adversarial, if they're sort of your enemies in some ways or your opponents in some competition for a job, being exposed to them might things, in fact, if anything, worse.

Finally, there's another piece of very nice evidence by Corno et al. This paper is considering the impact of random interracial interactions among college roommates in South Africa on stereotypes, attitudes, and performance. So what they look at, essentially, is roommates of different race reduces-- so these are black and white students-- reduces white students stereotypes towards blacks and increases interracial friendships.

It also improves grades and lowers dropouts among blacks. So there's sort of a number of positive effects of a very simple policy of increasing contacts among roommates. And again, if you think about this evidence, it seems to suggest that perhaps that's kind of like if you have a roommate that you're going to live with for quite a while, that feels a lot like collaborative contact in the sense of, like, you're sort of stuck with that roommate for quite a while, so you might as well sort make the best out of it, even if you initially don't like that person and you have the same objectives of being happy together, living together. And that really seems to foster and have some positive benefits.

So taking together, that sort of says, so A, the contact hypothesis seems to be broadly right. Contact to people who are different from you might make you more tolerant or more prosocial towards them. There might be more sort of integration of different groups. But what really seems to matter quite a bit is the type of contact that people are exposed to.

So finally and relatively quickly, I'll tell you a little bit about whether people underestimate the benefits of prosociology. I should mention that problem set 3 question 2 is, in fact, trying to ask you to do something related to that. So you might want to actually do the problems at first before you finish this lecture. At least sort of-- sorry, not the entire problem set, but question 2 of problems set 3. That will not take you very long to do so, and it's more like a fun exercise that I thought would be nice for you to engage in.

But anyway, this is a very nice paper by Kumar and Epley. That's a typo here. It should say Kumar and Epley. And the officer is asked the question about, do we have correct beliefs about the impacts of generosity? And so what's the underlying reason here is, well, many prosocial acts require estimating the impacts on the recipient.

If you give money to somebody in Africa, if you help anybody and people in need, if you donate money in general, if you write letters of gratitude or if you do random acts of kindness, it requires some sort of estimation of, how is the other person going to feel if they receive-- if they are on the receiving end of this prosocial act?

Now, what Epley and Kumar argue is people are subject to egocentric bias, which may lead them to systematically underestimate the positive impact of prosociology. In this case, gratitude letters. And so why is that? Well, it's because predicting others mental states is difficult. It's really hard for people to understand how another person might feel.

Usually, people would sort think about themselves and sort think about how would I feel, and what would things be, and how are things like. And it's very hard to understand how others might react to you. It sort of requires perspective taking, and that's sometimes tricky for people to do.

So this is again a very nice experiment by Kumar and Nicholas Epley. Not Nicholas last name, but Nicholas Epley. And they test whether people misunderstand the consequences of showing appreciation. And so what are these experiments? There's a series of experiments that look at this. What do they look like?

But what they do is they ask people, MBA students or subjects in experiments, to pick a prosocial act. That's writing a gratitude letter, for example. And then ask how the giver and the recipient will be affected. In some cases, it's only the recipient, and in some cases, also the giver.

Then they perform the prosocial act by, for example, writing a letter of gratitude. And then assess how the giver under recipients were actually affected by doing that. And then you can compare that to item number 2. To the estimation *x ante*.

Now, what do they find? They find clear evidence that when you look at, like, in particular, when you look at the giver-- so these graphs all show kind of the predicted ratings. The ratings is between 0 and 10 about what's the experience? How happy is the other person? What is the surprise about receiving this letter of gratitude in this example?

So people are more surprised than predicted. They are also more surprised about the content than predicted. They're also-- the recipients mood is actually better than predicted. And the awkwardness is also better than predicted.

So when you ask people to write letters of gratitude, people tend to say, oh, you know, it's going to be tedious to do, and the other person might feel it's awkward. And what am I going to say? And is it going to be weird when I'm going to say? Is the person-- don't they know already anyway that I'm really appreciate of that person? And so on.

So people come up with all sorts of reasons why they might not want to do that. If you sort of introspect and ask yourself how many letters of gratitude have you written in the last year, that doesn't-- few people actually do that in practice. And perhaps some of these reasons like the perceived awkwardness or perhaps sort of the underestimation of the recipient's mood might contribute to that.

Now, if you also think for a little bit, like, how would one actually feel to receive these letters, it's pretty obvious once you think about it that most people are actually quite happy about receiving such letters. It's kind of nice if somebody tells you, look. You did something really nice for me some time a long time ago, and this helped me a lot in my career in whatever way, or getting into school, into college or whatever it might be. It's really nice to sort of hear from somebody that you did something nice in their life that people are quite happy about.

Now, what's sort of some evidence, some summary of that? Well, so we tend to systematically underestimate other's appreciated and expressions of gratitude. That's also true for some other types of effect. That tends to be also true for random acts of kindness. People tend to be surprisingly more happy about those compared to predicted.

It's also true-- so Epley and Schroeder argue for social connections. So people also underestimate on average how they themselves and others feel when starting conversations. So what this experiment does is this gets people to-- it randomized people to start conversations during commuting on buses or trains. And once people start doing that-- again, before, when you ask them, like, how is the other person going to feel, how are you going to feel, and so on, people will say, well, it's going to be awkward. And what are we going to talk about? And is the other person even interested? And so on. But when you actually do that, people seem to be quite happy to have started conversations and making human connection.

To be clear, not everybody is happy. Some people might be also grumpy. But the vast majority, at least in these types of study, seem to be quite happy about initiating social contact, about expressions, letters of gratitude, or things like random acts of kindness.

It's really nice if somebody happens to be something to do something nice towards you. Could be like a random person who you've never seen before and you never see again. Just some random person on the street is really nice to you. That might make your day. It could be also somebody who quite well, and he's a good friend, and who just wants to do something nice for you for no good, apparent reasons.

So one caveat to these kinds of experiments is these are very much like one-shot, short interreactions. These are one-time interactions, and then the effects are measured. So there are some questions about, do these effects persist for repeat interactions?

That is to say, like, maybe if you do that once, people are quite happy. But when you try it more often, these effects tend to go away. So there's a bit of a question kind of like, how persistent, how important are these in practice if you do that more often in particular in the long run?

And then another important question here is then the question about always under-investing in prosociology in terms of our behavior. So lots of people who tend to be quite selfish and do good things for themselves and not so much others-- presumably, they do that because they want to make themselves happy. Presumably, people who do nice things towards others to some degree do that because they want to make the other person happy. In part perhaps because of social image and self-image concerns.

But one important hypothesis that Epley and others raises the question of, like, well, are we underestimating how good it might not only make others feel but also ourselves feel from being nice? That is to say, perhaps, one easy way of making ourselves happy is not just by being selfish and maximizing whatever outcomes but really, being nice towards others, in part perhaps because it just makes us happy to see when others are happy.

Perhaps it makes us happy because others will be then nicer to us. And there's a question about, can we make others and ourselves happier by being more prosocial, perhaps because we underestimate to start with what these effects might be? I don't think there's that much evidence on this specific question, but I'd love to learn more.

And when you think about your own life, you might want to experiment for a while and seeing trying to be nice or trying to engage in random acts of kindness, letters of gratitude and so on, that might be a nice habit to acquire. And you might see it may or might not make you happier. Surely, it will make the other person on the receiving end happier.

OK. So let me sort of summarize what we learned or what we studied on social preferences. So first, others' outcomes and utility matter for people's choices quite a bit. So in various situations, essentially, people are willing to give to others, and they're influenced by others in their choices.

Now, upon closer look, there's not much evidence of pure altruism. Rarely does it seem that people just do stuff for others just for the sake of others doing well. That is to say, if people get the chance of hiding they're not so nice actions, they will often take advantage of that.

The motivation there often is instead motivation to give to others or be nice to others to be prosocial is often saving face in front of others-- this is social image-- or themselves, which is self-image. So here-- and because of that, situational circumstances matter greatly. Societal norms are really important to consider. So when you think about incentives or any sort of types of structures and organizations, how might you-- incentives, or how might be sort of a structure certain groups of people working together, understanding these norms and circumstances is key for fostering prosociology.

If you want people to be nice to each other, you have to set it up in a way that it's maybe observable what people do. It's encouraged. There's opportunities for reciprocity, and so on. So a lot of the design of society, of a firm, of a group that you work in or a team that you work in really seems to matter.

So on the one hand, as I said before, it's a little bit disappointing that there's not much pure altruism. But on the other hand, if you sort of understand the motivations for people being to engage in acts that are good for others, you can sort of design incentives and structures that people work in accordingly, which will then create people being friendly and cooperative to each other.

Second, I showed you that social preferences matter at workplaces. Relative pay can depress incentives to work. This is evidenced by Bandiera. The fruit pickers that I showed you. Pay inequality can lower performance via reduced morale. So in particular, if pay inequality is seemingly unjustified, that's really bad for worker morale and might lower our worker outputs and sort of attendance and so on.

Third, social preferences appear to be malleable and shaped by external factors. In particular, there's evidence in favor of the contact hypothesis. Being exposed to others from different backgrounds really seems to make us more tolerant and more understanding, more prosocial towards these other groups. But also more prosocial perhaps in general, which is the evidence by Gautam Rao.

And then finally, there's some evidence that biased beliefs may lower prosociology in the sense that people might under-invest potentially in how nice they are to others perhaps because they misunderstand what the effective of engaging on such a prosocial act might have on others. And so potentially correcting these beliefs or experimenting might sort of increase prosociology, at least in some settings.

Now, what's coming next? What are the next lectures? Lecture number 14 is about limited attention. 15 is about projection and attribution bias. This is the idea that people have trouble projecting how they might feel on different states of the world. That is the idea if you're really hungry, it's really hard to project how you might feel when you're not hungry and vice versa. We're going to look at that.

In lecture 16 and 17, we'll look more specifically about beliefs and learning. We talked a little bit about beliefs already in various ways. But now, we're going to talk specifically about biases in beliefs and learning, as in people sub-optimally learning perhaps because of computational issues in the sense that it's really hard to learn in some settings, in part because of motivated beliefs.

People might drive utility from their beliefs. As in, for example, I might want to think that I'm a good-looking, smart, and a good teacher. And if I get feedback on one way or the other, I might react to positive feedback a lot more. I might sort of update positively if somebody says, Frank, you're really smart. I might sort of update positively because I feel really good about it. If somebody says instead, Frank, you're not so smart, I might mostly ignore that feedback because it might make me feel bad about myself. So that's all for now. Thank you so much.