PRINCETON UNIVERSITY
BENDHEIM CENTER
FOR FINANCE

markus'academy

Raj Chetty on Social Capital and Academic Mobility

September 8, 2022

Webinar Transcript

Transcript:

Markus Brunnermeier: Welcome back everyone to another webinar organized by Princeton, we are very happy to have Raj Chetty with us from Harvard University today, hi Raj,

Raj Chetty: Hi Markus, great to be here.

Markus Brunnermeier: Raj will talk about social capital and economic mobility, and some recent work he did in Nature and other journals together with his group at Harvard and others. So let me put things in perspective briefly, and then we move forward to Raj. So first I would like to relate social mobility to inequality, so there are different inequalities we are studying. In economics we are studying income inequality which is like a flow measure, or wealth inequality which is more of a snapshot of where people are across society relative to each other. And then there is social mobility which is more of a dynamic measure, it is about how do wealthy people become wealthy or the poor become rich or the rich become poor, how big is the change over time, so it is a more dynamic perspective. So it is a switch from poor to rich and vice versa. The question is can you have upward mobility all the time or is it the case which when someone moves up, someone else is to move down, and that is essentially one thing that people can have is one thing over time, with higher inequality you start out with huge inequality and over time it will go down, so it will go that everything is moving up to a point, but it could also be that there are immigrants coming from outside of the country who start poor and become wealthier and become wealthy as those initially in the population already. But you can have upward mobility but typically a lot of social mobility means not only upward but also downward social mobility as well. But what I would like to introduce is one of the ideas in my book, the Resilient Society, is resilience inequality, so people differ not only in their income or wealth or other status, they may also differ in how resilient they are, so some people might bounce back very easily from shocks while others might not bounce back very easily from shocks, so even if they are equally wealthy, have the same income up front, they have actually different chances to take opportunities. If you easily bounce back from negative shocks, you can take on some opportunities, and then bounce back from it. Someone who is not able to bounce back is scared to take on opportunities, and that leads them subsequently to income inequality, and it builds up as people get more income over time. They become wealthier, and this leads to wealth inequality. So resilience inequality is also something which I am very fond of studying. My second slide is to say what can we say about social mobility where there is resilience. Social mobility is like the poor steadily moving up if it is upward mobility. And the question is if there is no ceiling for people to move up, that is very important. Resilience is very closely related to that, but it is always conditional after having a negative shock, so once you have a negative shock, can you bounce back? And social capital is the network, the community, or facebook friends, is this the case where you can bounce back, so it is not, in general, unconditional. It is more conditional on a negative shock, and probably one can use some of this data to study resilience, or the question can be studied in a resilience mindset itself, once you suffer a negative shock can you bounce back? And once you have a lot of friends who can help you to bounce back, you can also take more opportunities.

3:45

Finally I would like to relate it to US immigrants, and that relates to an earlier webinar we have with Leah Boustan from Princeton, she studied US immigrants over many, many decades and centuries and she finds that immigrants are upward mobile, as they were in the early 20th century, so there is nothing worse than that, and she also looks very much across generations, and I was wondering whether the Facebook data matches, whether you can also identify immigrants, and whether you find differences between immigrants and non-immigrants. Perhaps we will already see or it is already there, and we will see more about that. So let me go to the poll questions Raj put forward, and then we will come to that. So the first question was "which type of social capital is more strongly associated with economic mobility?" Is it interaction across class lines, so the degree of interaction across class lines is about 76%. The rate of volunteering is only 4%. The density of civic organizations, 14%. And the extent of "cliques" is only 6%. So a huge majority is the degree of interactions across class lines, so that's what people thought. Perhaps they have already read the articles Raj has published. Second question was how highly connected networks help low-income children to escape poverty, so pick one of the mechanisms you think is the most important one, is it shaping aspirations? Is it providing access to information? Or is it providing professional references? And here the answer was a little bit more evenly distributed. 39% for shaping the aspirations, 44% for providing access to information, and only 17% said referrals or professional references are really important. Third question was about why do lower income people have fewer high income friends, or which factors are the most important ones? The first lack of exposure, which is like segregation, is about 30%. Friend bias, which is a lack of interaction, conditional on exposure, so given that you have exposure, you might still have to go to the same college, but you still have no interaction, that is 16%. Both matter is 54%. So 30%, 16%, 54%. And finally, the last question is where to befriend a peer from a different social class, so in which settings? Workplace is 10%, highschool is 35%, college is 22% and religious groups is 33%, so by far the biggest is high schools and religious groups, around a third each, and then college is 22% and workplace is only 10%. So that is the prior belief, everyone goes into it and then we will see what we will learn and hopefully get a different picture or perception on Raj's findings. Thanks again for doing it and looking forward to your presentation.

7:14

Raj Chetty: Great, thank you so much Markus, it is a pleasure to be back with you, and thank you everyone for joining. So I am going to share my screen here, toggle over to these slides, so I am going to talk about the work on social capital and economic mobility, with the very nice introduction that Markus gave to these issues, and I am going to be presenting some work that, joint with a large team of folks who have been working on these issues over the last 4 years, and we just released these finals and the paper was published in Nature last month, and I am going to describe them here. So to motivate what I am going to talk about today, I want to set the stage by talking about some earlier work that I did with my colleagues, Nathaniel Hendren, Maggie R. Jones & Sonya R. Porter, where we look at teh geography of upward mobility in the US, so before we look at the data we use in this paper, in earlier work we used anonymized information from tax returns covering the entire US population, linked to census data, in order to measure a child's chance of rising up in the income distribution and achieving the "American Dream" so to speak, varied across the United States. So let me describe how we constructed this map, and what we have learned from it, so what we did is we took data on all kids born in the US in the early 1980s, about 20 million children, mapped them back to where they grew up, and for the subset of kids who grew up in low income families, that is families that were earning about 27 thousand dollars a year, which puts them at about the 25th percentile of the national income distribution, we asked them what is the average income percentile that these kids reach when they themselves become adults, where we are measuring their incomes using tax data at age 35. We color the map so that blue/green colors represent areas with higher levels of upward mobility, where kids from these areas are more likely to rise up, and red/orange are places with lower levels of upward mobility. So if you look at the map, you can see and draw geographic patterns for yourself. THe center of the country and parts of the coast look great in upward mobility if you start in a family in the 25th percentile, you have a good chance on average of being in above the 50th percentile, so high social mobility in one generation. But if you look at other parts of the country like much of the southeast for example vs like detroit and cleveland and the rest of the midwest, kids who grew up in families of the exact same income level have much poorer prospects of rising up, and many of them are not reaching higher levels of income, and the parents on average in many of these places are below the 35th percentile, showing the variance across areas. This has captivated the interest of many scholars in the last few years, because it provides an unprecedented lens to study the determinants of social mobility and economic opportunity. Because we can essentially ask you what is different in some of these counties, in the rural midwest, for example, versus the rural southeast, or one city versus another city, or even zooming down to the census tract level, and ask, what are the drivers of differences in economic opportunity? And so there's a large literature that has emerged over the past few years that has identified several strong predictors of this variation in upward mobility that we're seeing across areas in the map that I just showed you. And it's identified a variety of factors that seem quite important in predicting these differences in the economy. So an example: a number of studies have shown that places with lower poverty rates, more mixed income areas tend to be places where kids from poor families are more likely to rise up. There are strong correlations with things like school quality income inequality, so linked to what Markus was talking about at the beginning on different concepts of inequality. There's some well-known work going back to Myles Korak and Alan Krueger, who was at Princeton, showing that there's a strong correlation between levels of income inequality within a generation, and rates of social mobility across generations. More unequal societies tend to have less social mobility as well across generations. There are well-documented links with racial segregation family structure and a number of other factors. Now, in the course of following this literature, and contributing to it myself, a number of people mentioned to me, and I myself thought you know it could be quite intuitive that the idea of social capital might be important in driving some of these differences in economic mobility. So my colleague here at Harvard, Bob Putnam, has written a series of books starting with Bowling Alone, which many of you might know, going back to the 1990s on the potential importance of social capital for a variety of outcomes, and there's been a fair bit of literature in sociology and other adjacent fields suggesting that the strength of our networks, or who we are connected to, might be important in our communities, in driving things like upward economic mobility. So motivated by that, I began talking with some colleagues in particular, Matt Jackson, who's a pioneer in the study of networks in economics at Stanford, as well as Theresa Klucher and Joanna Strobel, who've done some of the most important empirical work on networks in recent years in economics, and we teamed up with a group of others, including the Facebook core data science team, to think about whether we might be able to quantify social capital systematically use sort of a big data approach to understand what exactly social capital is, measure it systematically, and try to understand if and how it's related to economic mobility, and how we might be able to manipulate it through policy going forward.

13:02

Raj Chetty: And so, before we get into that work of measurement and empirical analysis, the first step for us was to actually try to think in a precise way about what social capital actually is. If we're going to try to measure something, the first step is to try to define it precisely. Now, this is actually more challenging than it might seem, because there are many different concepts of social capital that have been used in the prior literature. People sometimes mean different things by this term. So what we do is organize these many concepts that people have proposed into three broad categories. The first is what I'm going to call measures of connectedness, the extent of which different types of people are connected to each other. So a leading example might be the extent to which low-income and high-income people in a community are friends with each other. We're going to call that economic connectedness. But you could also think about connectedness along other lines, for instance, by age, or by language, or by race, et cetera. Second set of concepts is what we call cohesiveness. Think of this as the extent to which your network is fragmented to separate cliques. To what extent is it tight knit? Are your friends that turn friends with each other? Mathematically, the way to think about these concepts is basically strip away the labels or the colors from the depiction on the left and just to ask, What does the network graph itself look like? One famous concept introduced by the University of Chicago sociologist, James Coleman, is the idea of closure or clustering in networks. If you have two friends, what is the probability that they are in turn friends with each other? How many triangles are there in networks? And so that's another concept of social capital. You know, basically measuring how tight knit a community is which could potentially be relevant, as Coleman argued for a variety of outcomes. A third and final category that we'll look at, which is completely distinct from the network-based measures, are what you might think of as measures of civic engagement. So these are some of the measures that Bob Putman popularized things like rates of volunteering or participation in civic organizations. How strong is the community? So you can think of these last set of measures as not using network data. The cohesiveness measures only use network data, and then the connectedness measures combine network data with data on people's characteristics. So that's kind of a partition of the space of potential measures that you might think about, and we will measure each of these. Our goal is to measure each of these systematically and understand if and how they're related to economic mobility. So the way I'm going to organize our discussion in the coming few minutes is to talk about what we did in four steps in these two papers starting with how we measure the concepts of social capital that I just laid out, using very rich data from Facebook through our collaboration with their internal teams. I'll then analyze the associations between these new measures of social capital that we construct and make publicly available with economic mobility, having shown that one particular form of social capital, economic connectedness, is very strongly related to economic mobility. I'll then turn to identifying the determinants of social connections. What is it that makes people have more social capital in some places relative to others? And then, finally, I'll take a few minutes to talk about the data that we've released as part of this project that we think will hopefully be helpful in targeting interventions and moving the field forward in the future.

16:28

So let me dive in by talking about how we measure social capital using these new data, starting by just giving you a quick overview of the underlying micro data that we use to construct these measures of social capital. So we're going to use data from Facebook in particular. All US-based Facebook users between the ages of 25-44 who are active on the platform have at least one hundred friends in the US. And we're going to take a snapshot of that data for the final analysis. We're showing you here as of May 28, 2022. That comprises about seventy two million people that sample between them. They have about twenty one billion friendships. And importantly, the reason we focus on that age range is that's an age range where virtually everybody in that age group is using the Facebook platform. The seventy two million people represented there cover eighty four percent of the US population between ages 25 and 44. So we're starting from a pretty good place in terms of a representation of the population. I'll talk a little bit. Ah going forward about that remaining sixteen. Should we worry about selection issues and so forth, in terms of who's using the Facebook platform. But the point here is we're starting from a good place.

Markus Brunnermeier: So Raj, someone ask whether you also look at the family structures as well, and closeness of family, to have Facebook friends within the family as well or so.

Raj Chetty: Yes, so the bulk of the friends we're gonna look at are going to be outside the family, because, as you will see, people have hundreds of friends on Facebook, and that's gonna be predominantly focused. We haven't delved into it in a great detail. But you can look at familiar relationship relations as well within the data. I'll show you a little bit on this when I show you a cut. Looking at closer friends versus friends who might be more distant and obviously familiar links might be closest links. And let me come back to that. There we find fairly similar patterns in those two cases. But I think there's more one could do for the family going forward as well.

Markus Brunnermeier: and another quick question from Nairobi. He would like to know whether you also take into account how much time they spend on Facebook. Is it just connected? Friends?

Raj Chetty: Yeah. So in our basic analysis we're just going to look at who your friends are. You can then look at measures of intensity of friendship. Time being one example, our time spent on the platform, and, as I will show you, we end up finding quite similar patterns. When we look at those stronger links and let me come back to comment on why I think that is. Great. So with that let me now dive into the first measure that we're going to construct, which is this measure of economic connectedness. This is basically a measure of the extent to which people are interacting across class lines. To what extent are people from low income or low socioeconomic status backgrounds interacting with people from high socioeconomic status backgrounds.

19:15

Raj Chetty: So why do we start with this measure? You know it doesn't come out of thin air. There are many reasons that people have speculated in prior work, and I've shown some evidence that economic connectedness of this type might matter for outcomes. You know, what are some mechanisms? You might think of things like information. So maybe if you're linked to somebody who went to college. They give you information about how to apply to college, how to prepare for knowledge, and so forth. It might have an influence on your aspirations and preferences. If you see somebody succeeding following a particular career pathway, maybe you're inspired to pursue that career in science or technology or business yourself, and then maybe there's a direct effect on job referrals. Lots of jobs in the US and other countries are obtained for a referral, and if you're in a network where you're connected to people who are higher income and have access to better jobs, maybe that directly impacts your economic opportunity. So there are lots of reasons to think this might matter. We're not going to directly unpack these mechanisms in this work. We're just going to be motivated by that and try to assess, you know, does economic connectedness actually relate to economic growth? So with that motivation, let me talk briefly about how we're actually going to measure economic connectedness. And there's a key ingredient we need. In addition to knowing your friendship network, we need to actually know your socioeconomic status right? So the way we measure socioeconomic status in the Facebook data is by combining several proxies, things like your zip code, because where you live tends to be very strongly predictive of your income given the amount of residential segregation in the US; the college that you went to, which many people self-report on the Facebook platform; things like the price of your phone model. If you have the latest iphone, you probably have a higher income on average than if you have an older model, et cetera. So we essentially take a combination of the set of measures, use a machine learning model to construct the best predictor of median household income in a person's block group, which is available for a subset of users on the Facebook platform. And essentially think of this as we're using this machine learning model in particular, our gradient boosted regression tree to combine these twenty different variables into a single index of socioeconomic status. Okay, we then take that combined measure, and rank people in the national distribution, relative to others in their birth cohort, based on their predicted socioeconomic status rank, so essentially one to one hundred in terms of percentiles, comparing you to other people of the same age. Now, these socioeconomic status measures that we've constructed, you know, you might worry, and we spent a lot of time worrying about whether they are accurate measures of socioeconomic status. And so we did a lot of benchmarking exercises where we show that these measures are very highly correlated with publicly available statistics, from administrative data sources on income distributions by high school, by college, by zip code. What we see in the Facebook data on average for each of those units, if we look high school by high school in America, for example, lines up very closely with what you would get from the National Center for Education Statistics, for instance. One final note on this we also obtain pretty similar results when we use simpler measures of socio-economic status, like just the median income in your zip code or simple, weighted average of standardized proxies rather than this fancier machine learning model. So the details turn out not to be that important in the end here. But we, you know, I just want to be clear about how we're measuring socioeconomic status combining all these different processes. Okay.

22:40

Raj Chetty: So using those measures, combined with the friendship network data from Facebook, I can show you this first chart here, which is a depiction of the degree of homophily the degree of cross-class interaction in the United States as a whole. So what this is plotting is for each percentile of the socioeconomic status distribution on the x-axis. We're asking, what is the average socio-economic status, rank, or average income rank which is a shorthand of your friends, and what you can see is that there's a very clear, upward-sloping relationship. The richer you are, the richer are your friends. That's a pattern that's been documented in prior work and we see that extremely clearly in these data as well. Now, some of the questions that were raised initially are essentially asking. You know, if you look at closer friends, people, you're spending more time with people you're more closely connected to. What do the patterns look like? I think that's an important question, especially in the context of the Facebook data, because the Median person has something like five hundred friends on the Facebook platform. And so you might have the reaction: Well, you know I have five hundred friends on Facebook, but they aren't really my friends, like the people who might influence my economic mobility. The people who really matter, say, five or ten people who run very close with. And so what we do to assess that is, look at your top ten friends defined based on things like intensity of interaction and how close you are, how many messages you're exchanging, etc. And what we find is, if you replicate this graph for your top ten friends, you find a little bit more homophily than you do overall, but the pattern looks extremely similar, and more generally, Markus, you know, coming back to those questions, if I replicate each of the results that I'm going to show you going forward within this top ten friend sample, you get extremely similar results to what I'm showing you.

Markus Brunnermeier: So I was wondering if you control for regions, if somebody lives in New York City, and all his friends will be wealthy, too, or similar in San Francisco, and then other regions in the US, you still find it?

Raj Chetty: Yes, this is why you still find it. So I'm going to get into that in great detail. That's basically going to be the distinction between exposure and friending bias that I'm going to talk about. So exposure is just. You tend to be friends with the people near you, as you just suggested. If you live in the area. You might have more high-income friends that's part of what's driving this. But there's another phenomenon, which is, even if I live near a wealthy person, I may not become friends with that person, or what do you call that friending bias? And I'm going to get there in a few minutes. How much of this relationship is driven by the first factor versus the second.

Markus: So one is our quick question on Yes, please, how to deal with duplicates and Facebook accounts.

Raj Chetty: Yeah, So there's a pretty um, you know. Sophisticated procedure that Facebook tries to use to eliminate duplicates using things like where people are logging in from et cetera. And we are taking advantage of that here now. Is it absolutely perfect that there's literally no duplicates? You know, I think one can’t be entirely sure, but our sense is that for the most part, you can think of these as new users.

25:38

Okay, one final benchmarking exercise any time we use a novel data set like this, you know. I like to make sure, using comparisons to other data sources in many different ways, to validate the new data. And so one validation that I think is very useful is to compare the degree of homophily that we're seeing in the Facebook data shown here in the green to what we see in a representative survey, the Ad Health Survey, which surveys about twelve thousand high school kids and asks them about their friends in high school, and then constructs you can construct, using that data measures of homophily by parental socioeconomic status. And what we're doing here is comparing what we get in Facebook to what we get in the ad health data. And what you can see is, I think, exactly what you'd expect if you think of the Facebook data as basically a national population estimate using seventy two million users. And then here's an estimate using twelve thousand users. It's basically like a noisy version of what you're seeing in the green. So the fact that those two align so well, I think, should be quite reassuring in terms of the use of this data in understanding these kinds of questions.

Markus Brunnermeier: Are the green dots so smooth?

Raj Chetty: So you know, part of it is just a simple consequence of averaging. So this is the conditional expectation. And what this is basically telling you is the conditional expectation function is very smooth, and there's of course, a lot of variance within each of those dots. And in a sense that's what I'm going to get to next, starting to unpack that variation. But you're right to observe, Markus, and I mean that's the power of big data, right that I think with enough averaging. We start to see these very clear, smooth relationships in populations. So you know It's a good segue now to then this aggregation to trying to understand what's behind that curve. The power of having such a large sample is not just that you get a prettier rough, and what I just showed you, you know nationally, but that you can now cut that draft in many different ways to start to understand the determinants of social capital, and how they affect various things. And so, in particular, we can disaggregate by zip code, by high school or college. So, in order to do that, rather than showing you that curve again and again for every zip code, for every high school, you know that would be a lot to look at. I'm going to distill it to a one-dimensional statistic that captures the extent of cross- interaction, and so in particular. I'm going to look at a very intuitive measure, what is the number of high SES friends, above median socioeconomic status friends that below median socioeconomic status people have as a share of their total number of friends. So this is just saying: what fraction of your friends are above median income if you are a below median income person, and we're going to take that measure and normalize it, we're going to divide it by a half, recognizing that if you main friends at random, fifty percent of your friends would be above median income if you're a below Median income person. So if this number is one. It means that you're making friends essentially at random in the population, and if that number is below one, it means that you have fewer high SES friends than you would if there were no homophily in the population. Okay, So this E.C. statistic is what I'm going to focus on throughout much of the presentation overall in the nation as a whole. That number is point seven eight, reflecting exactly the pattern I showed you on the previous slide that low-income people tend to be friends with lower income folks more than higher income folks and so that .78 number. The way to interpret that is that we have about a twenty, two percent underrepresentation of high-income friends relative to the random friending benchmark. That's the amount of social disconnection that we're seeing in the US as a whole.

29:20

Raj Chetty: Okay. So now, what we can do is construct that same statistic: what is the degree of economic connectedness of low-income people to high-income for people for every county in America, and that's what's being shown on this map, where we're plotting that EC measure. You can see it's centered around .78, which is the national average. Red colors now depict places with lower economic connectedness for the poorer, more disconnected from the rich and blue colors. You can see the numbers get above one in some of these places. No more than half of your friends are high income, right? So you can see in the map that there's a broad, you know. There's quite a substantial amount of variation in the degree of cross- interaction across places, as Markus was anticipating. Some of that comes from the simple fact that if you live in a richer place, you're going to interact more with high-income folks. But some of it doesn't come from that, as I'll get to in a second. We're going to see that it's about fifty over fifty in terms of those two explanations. Another point which you're probably already noticing, and I'm going to get into in much more detail in a second is that this map looks visually very similar to the first map that I opened with on economic mobility. And so we'll see in a second that these two things are indeed quite tightly linked.

Markus Brunnermeier: So one one quick clarification. When you talk about median income, you mean the median in your county or the median nationally?

Raj Chetty: It's always nationally. So we're ranking everyone nationally. And then I'm going to basically look at how much of the variation is coming from exposure to higher income people in your county or in years of code in your area versus in differences in interaction between those people. But all ranks are always in the national distribution. Thank you for clarifying. So this is a county-level picture in the nation as a whole. I just want to show you that a lot of this variation emerges not just across broad regions of the country, like the midwest versus the southeast, but actually much more locally. And so, in order to show you that, I'm going to toggle over to this tool here that you can access on the web yourself at a website, socialcapital.org, we're calling the social capital Atlas, where i'm going to start out with the national view that we saw in the previous map. But then, the way this tool works is you can type in any place you want. I'm going to type in New York City, given where Markus is, and just zoom in to look at the data. Now, zip code by zip goes to New York. The same statistic, What fraction of the friends of low-income people are high income and you can see that there's a tremendous amount of variation, even, you know, very locally. So in Manhattan, for example, the lower income folks who live there are more likely to be connected to I for folks less so in the Bronx, less so in Brooklyn, more so in Queens relative to other places. And then you see variation in Newark versus other parts of New Jersey and so forth. So there's a lot of very local variation. Now, I've focused here in depth on one of the measures of social capital that connectedness measures. I want to briefly mention the other measures as well.

32:16

Raj Chetty: So a second concept that I mentioned at the beginning is this idea of cohesiveness, forget about people's incomes, and just ask the extent to which networks are clustered versus not, so how many triangles are there in the network? How tight knit is the community? Purple here are places that are less tight-knit, where you have two friends, and they tend not to be friends with each other, and green are places that are more tight-net where clustering is higher. And you see here what I think is a very intuitive but different spatial pattern from what we saw with economic connectedness. The more urban areas, which was exactly what James Coleman had anticipated thirty years ago, which have more transient populations, tend to have less clustered, less tight knit communities. The more suburban areas out here in Long Island parts of suburban New Jersey tend to have much more tight-knit communities. The key point I want to make here for our purposes is this measure of social capital, especially the patterns that look very different from economic connectedness. It's not that Manhattan is high in that form of social capital, and the Bronx is lower, et cetera. It's a totally different spatial pattern. And then, finally, I'll show you a third measure of civic engagement. So here, with the Facebook data, we're able to measure in a much more precise way rates of volunteering than people have been able to do in the past. So in Facebook people join various types of groups, as you might know, and we can classify some of those groups as being about volunteering, and use that to construct a measure of the rate at which we are volunteering zip code by zip code in the US. And again we validate this at broader geographies, using data from a general social survey, and it winds up very closely with other measures of volunteering. And once again you see a spatial pattern that looks very different from what you see from the other measures of social capital where you know parts of Manhattan you see higher levels of volunteering relative to other parts, et cetera. And so the key point I want to make: there's a lot more that can be unpacked with these data, and this is all publicly available for download, and we hope people will explore this in greater detail. The simple point I want to make here is that these different measures of social capital each have very different spatial patterns. It's not like there is one place that's rich in all forms of social capital that people have talked about, and poor, and all others. It really matters which dimension you're talking about. And so, just to formalize that a little bit. Let me come back to the slides and just show you a correlation matrix of these different social capital measures that we've released publicly: measures of connectedness across different characteristics, measures of cohesiveness, measures of civic engagement. And when you look at the correlations between these measures, you can see there are some correlations that are higher than others, but generally they are pretty small. And so the key thing, I think that underscores Markus, is that when we think about social capital we need to be precise about what type of social capital we're actually talking about, just rather than talking about it in broad terms. And so with that motivation. I'm. Now going to turn to the next part of the talk and ask our motivating question: Are any of these measures related to social capital? And if so, which one? And so I'm going to start with economic connectedness where I'd already previewed the result. Here's now a scatterplot version of that where we're plotting rates of economic mobility from the prior work we did, using tax records for the two hundred largest counties in the US versus this new economic connectedness measure from the Facebook data, and you can see those things are incredibly, strongly linked with a correlation of about .65. Places where we're seeing low-income people have more high-income friends tend to be places where you have higher levels of economic mobility. So that's the correlation with economic connectedness shown as .65 here on this chart. Let's now correlate rates of upward mobility with all of these other measures of social capital that we've constructed. And you see, I think, an incredibly stark and clear pattern, which is that there is one and only one measure of social capital that's strongly correlated with mobility. And that's cross-class interaction, economic connectedness. All of these other measures have little or no correlation with economic mobility.

36:20

Raj Chetty: So that's a set of univariate correlations. You can do the same thing in a multivariable regression, put them all in on the right-hand side of a regression. Everything is explained by economic connectedness as opposed to these other measures. Right? So what we've established there is that economic connectedness is very strongly correlated with upward mobility, and it's possible that there's ah an important causal set of mechanisms there. I mentioned some of them before, things like aspirations, information, job referrals. But of course, economic connectedness may be correlated with mobility. Even in the absence of a causal effect for other reasons, and we need to consider those other reasons carefully to understand what's going on here. So I think there are three other channels that you might think about, for why we see a link between these two variables. The first is reverse causality. Perhaps it's that higher upward mobility leads to higher economic connectedness in adulthood rather than the other way around. What's a simple story that illustrates that? Suppose you start out as a low-income kid. You live in a place where people tend to rise up for other reasons. They have the schools, you know they have other good factors. You had a lot of lower-income friends while you were growing up through your family. Now you have a lot of high-income friends. Once you've gone to college and so forth, and that's going to lead to a high level of economic connectedness, I suppose not because connectedness, effective mobility, but rather the other way around. So that's one thing we'll have to think about. A second possibility is selection. Maybe it's just that the people who live in highly connected areas are different in other dimensions. Maybe they're from different racial backgrounds. We know from our prior work that there are big differences by race and economic mobility in the US. And so maybe it's about simple selection rather than connectedness actually matters. A third possibility is, even if we establish what places have a causal effect on upper mobility. Maybe it's not the direct effect of connectedness itself, but rather other features of neighborhoods. Things like the quality of the schools that might be higher in more economically connected places that generate higher levels of upward mobility. What I'm going to do in the next few minutes is walk through each of these other explanations and assess their importance. So let's start with reverse causality. So to address reverse causality. The simplest way to think about that is just timing logic. Let's look at a subset of friendships that are made before people enter the labor market based on parental SES. So essentially define the network based on a set of variables that are realized before you actually realize income in adulthood. Okay. So they cannot be mechanically affected by rates of upward mobility. So how are we going to construct these measures of childhood economic connectedness? We do it in two ways. First, for a subset of people in our Facebook analysis sample, we can link them to their parents, about thirty percent of people. We can find their parents in the Facebook data, and we can identify which high school they went to, and that we can construct measures of high school friendships based on parental socioeconomic status. So that's one approach.

39:24

Raj Chetty: A second approach is to use current day users of Instagram, which is now the most popular, one of the most popular platforms for people who are, say, between the ages of thirteen and eighteen today. And again, we can form a measure of childhood connections for that group, and it turns out when we use either of these measures of childhood economic connect, and as we find that they remain very strongly correlated with upward mobility. .44 in the facebook subsample link to parents. .62 using the Instagram full sample. So what that shows you is this can't just be purely about reverse causality. Even if I define your network purely based on your childhood friends, there's a strong link of upward mobility, so that, I think, deals with the reverse causality issue. Let me now come to the second issue of selection versus causal effects. So what we would like to do here is, look at the relationship between the causal effects of counties on upward mobility and economic connectedness. So just to step back and clarify ideally how one would do this, I find it useful to think about What's the ideal experiment you would run to purge selection? What you would do is randomly assign kids to different counties while growing up, and then test those who you assign to a higher, to a county where the poor are interacting more with the rich, whether they earn more as adults. Now, of course we can't run that experiment directly. In practice. We can't randomly assign people to all the counties in the Us. So what i'm going to do instead is make use of a quasi-experimental design that I developed in some earlier work with Nathan Hendren, and in a paper published in 2018 where we use a mover's design to identify the causal effect of each county, and rather than going to the details of that paper, I'm just going to give you a quick summary here of what we do: we look at seven million kids using data from tax returns whose parents moved to a different county while they were growing up, and we make the key identifying assumption that the age at which you moved between a given pair of areas was orthogonal to your potential outcomes conditional on parental and to sort of say that differently, if Markus and I both move, you know, from, say, New York to San Francisco. But Markus moved when he was four, and I moved when I was seven. We're basically going to compare that Markus's outcomes to my outcomes to identify the relative impact of growing up in San Francisco versus New York. We're going to make the identifying assumption that conditional on both of us making this, that same move, and having the same parental income, our outcomes would otherwise be be comparable, except for the fact that we grew up in different places, and using that, we can identify pairwise the effect of growing up in County A versus County B for every pair of counties in America, and then distill that to a vector of fixed effects, the vector of causal effects of each county immediately. So that approach, you know, there's an important identifying assumption there. It's been extensively tested and validated in the subsequent literature. I'm just going to take those estimates at face value here, and look at how they are related to economic connectedness. So the idea here, now that I'm putting the causal effect of growing up in a given county on the y-axis, constructed using this method that I just described, and plotting that in a bin scatter plot versus economic connectedness. This measure is now purged of selection under our identifying assumptions, and the key thing it shows you is that if you grow up for more years in a county that's more economically connected, you have a higher level of earnings in adulthood and the magnitude here, it's a little bit hard to interpret directly from this chart let me just say it's, you know, quite substantial. We estimate that if low-income kids were to grow up in counties with a level of connectedness that is comparable to the average high-income kid in the US, their earnings would rise by about twenty percent.

43:17

Okay, So there's quite a substantial connection between these two things, so that analysis in our view establishes that a large part of the correlation that we're seeing is related to a causal effect on upper mobility rather than selection effects in particular, moving to a higher connectedness area earlier childhood as a causal effect on upward mobility. So last step: what that shows us is that growing up in a higher, easy place is good for you in terms of your chances of rising up. But it doesn't directly establish that that is because of connectedness itself, as opposed to other factors that might be correlated with living in a more connected area. And Markus, I think, gave what I would think of as the leading example: more connected places tend to be richer places, places where there are more high-income people around. Those places are going to have more resources, better public goods so we tend to have property tax finance schools in the US. Your school is going to be better funded if you live in a more connected place on average, and maybe it's that factor that really matters for economic mobility rather than connectedness itself. So to get at that, the last thing we do in this part of the analysis is compare the explanatory power of the strongest predictors that have been identified in prior research, of differences in economic mobility across areas, things like poverty, rates, inequality of racial segregation, and so forth, versus economic connectedness. And I'm going to start by examining the role of average neighborhood incomes which is a very strong predictor, and importantly, is also highly relevant for policy. It's currently the most widely used marker of high opportunity areas. So if you look at a lot of policies in the US that are targeted at low versus high opportunity places, they often define what is a high opportunity place based on the level of income. So to understand the relative contribution of income versus economic connectedness, let me turn to this chart here. So what we're doing here is representing each zip code in America by a dot. And we're plotting economic connectedness and that's a code constructed using the Facebook data. The degree of cross-class interaction versus media and household income in the Zip code just from census data. You can see exactly as Markus was anticipating, there's a strong upward sloping relationship here. If you live in a richer place, you tend to have richer friends on average. But there's also a lot of dispersion around that. You can have a set of places that are comparably rich, where you see a lot more across class interaction than some of them relative to others. Now we get to what I see as the key point. Let's color these dots by the rates of upward mobility from the opportunity Atlas tax data. The earlier research that we had done. Remember, red colors are places with lower levels of economic mobility, where kids from low income families are less likely to rise up, and blue colors represent areas where kids from low income families are more likely to rise up. What you see is, I think, a very striking and important pattern, which is that If you take any vertical slice of this figure, there's a clear change in the color of the dots as you move up, so what that's saying is, if I take a set of places, all of which have a median household income of around fifty thousand dollars a year, and I move from a place where the poor are not interacting with the rich to a place where they are interacting with higher income. Folks, you see a systematic change in rates of upward mobility associated with that dimension.

46:38

Raj Chetty: If, on the contrary, I do the converse exercise and look at how rates of economic mobility change as I move to richer places relative to poorer places, so places with very high median household incomes versus low median household incomes, I see essentially no change in colors, so controlling for the level of cross-interaction resources seem to have essentially no relationship with economic mobility, so that analysis suggests– this is essentially a non-parametric depiction of a multivariable regression– what it's basically telling you is that it really seems like It's the economic connectedness dimension that matters not the level of resources itself. So we can do this type of analysis for a variety of other factors that people have identified in the literature. So one of the relationships that has received a lot of attention, both in academia and the policy conversation is the strong link between levels of income inequality, and upward mobility across generations, shown in this regression, here in the first column where we're addressing up for mobility on the Gini coefficient across counties. But then, if you go to the second column and include economic connectedness as an explanatory factor. You can see that that completely knocks out the relationship between inequality and mobility. In that sense, connectedness basically explains the relationship between inequality and mobility. So what is going on there is that more unequal areas tend to be places where there is more social disconnection between low and high-income folks, and that fully accounts in a statistical sense for why they have lower levels of economic growth. To give you another example, there's another quite well known strand of work which shows that more racially segregated areas, areas with larger black populations, for instance, tend to have lower levels of upward mobility for both black people and white people. Once again, if you look at this adjacent pair of columns, once we control for economic connectedness the relationship between share black and levels of economic mobility disappears completely. So this goes back to some well-known earlier work by my colleagues, David Cuddler, and Ed Glaser, in a paper in the QJE in 1997 which they documented, using other data. That segregation is extremely harmful for blacks, but they noted that we do not have an exact understanding of why that's true. And what we see here is that the lack of economic connectedness in racially segregated neighborhoods essentially provides a full statistical explanation of that pattern as well. So, um, you know what we've seen so far is that economic connectedness really seems like a very strong predictor of chances of rising up in the income distribution for kids from low-income families, and it explains a lot of the important associations found in prior work. And so that leads us to think that this is a very important factor to focus on going forward. Let me make one final point on the association of economic mobility which people often ask about. So we've seen that greater connectedness is associated with better outcomes for kids and low-income families. Does that come at the expense of outcomes for kids and high-income families? Is this kind of a zero-sum game, or can there be gains in terms? Of growth? This relates to some of the questions Markus was asking initially about whether upward mobility comes at the expense of greater downward mobility? And so to get at that, Let's plot the now not just the outcomes of low-income kids shown in the orange here in a bin scatterplot version that says the average outcomes income ranks in adulthood of kids and low-income families plotted against cross-class interaction. Let's now plot the same thing in the green for kids starting out in high-income families, and you can see that there is a downward-sloping relationship here, but it's much flatter than the one that you see for low-income kids, a much shallower slope. But, moreover, and I think this is really key, once we control for the share of high-income residents in an area, so control in a sense for the resources available in an area, the average level of incomes, that green relationship becomes basically flat. So there's no association between Cross-class interaction for high-income kids and ah, between cross-class interaction and upward mobility for for high-income kids, whereas there remains a strong relationship for low income kids, as we were seeing earlier, and so that's addressed that if we look at communities where there's more cross-class interaction holding fixed the level of resources, it might not actually be a zero-sum game, we might have a situation where you know, quite encouraging that we might have low-income kids rising up at higher rates without coming at the expense of hiring advocates, and I think you know well there's more work to be done on this. It's certainly suggested that this is not a mechanically a zero-sum situation.

51:16

Markus Brunnermeier: So Raj can I push you a little bit in a different direction in a sense that you say well, I'm connected to somebody above the median income. Could it be so much better to be connected with someone who is in the top 10%. Did you look at this too?

Raj Chetty: So the above the median split, Markus is just, you know, one way to sort of the data, and you know we we we do other splits by quintile, by decile, and so forth, and generally you will find, you know, correlated patterns, because the rate of above below median linkage is very highly correlated with the rate of, say, bottom quintile linkage. Now, one thing that I think is interesting, and we're hoping to explore in future work is the idea that there are very few connections typically to people who are very far away from you. One theory, as you were kind of suggesting, is that they might potentially be quite helpful if you manage to make them, but there are other interesting models out there. There's some nice work by Deborah Drey at NYU, which argues that if you think about models of how role models might work, if somebody is too far away from you, it might actually feel unattainable and be less inspiring, in a sense, than somebody who's a little bit ahead of you. So it's actually not totally clear to me that it's always better, even if you can make that connection to the person who's, you know, in the top one percent of the distribution

Markus Brunnermeier: And your data allows you to analyze that?

Raj Chetty: Exactly, and it's something we haven't done yet, but we can try to unpack going forward which of those connections seems most valuable. For now we're just saying, you know, on average, being connected to higher income folks seems good. Great. So in the remaining few minutes here, I want to take a little bit of time talking about the second paper, where, you know, having spent a lot of time establishing that this idea of cross-class interaction seems really important for economic mobility, we wanted to understand what the determinants of economic connectedness are. What is it that's making people interact more with income people in some communities relative to others? And what implications might that have for policy going forward if we want to create a more connected society, either to increase mobility or just because we want to live in a society that's less disconnected for other reasons. So to set up this work, it's first again, useful to start with a conceptual framework of what it is we're trying to get at. And here I think it's useful to split up the determinants of economic connectedness into two very different forces.

53:43

Raj Chetty: The first is a little bit of what you were getting at with your early questions, Markus, on what we're calling exposure, or just segregation by income. So take this example here, where we have two schools, school A and school B, and all the high-income kids go to the first school and all the low-income kids go to the second school. So in this kind of situation, where you have a lot of segregation, you're presumably going to have a lot of cross-class disconnection as well, because you can't really be friends with people you never meet. So that's one possible explanation for why we're seeing social disconnection in America. A very different explanation is that you actually have integrated schools, as in this example here. But still you don't have interaction between low and high income people. So a lack of interaction conditional on exposure. And we're going to call that friending bias. So it's critical, I think, to understand how much of the social disconnection we're seeing is due to a lack of exposure versus friending bias. Because the way you want to address the problem is very different. If it's the former as opposed to the latter right? If it's about the former, you want to think about things like changes in school district boundaries, or busing or affordable housing or basically ways to integrate people. If it's about the latter, you've got to think about what's going on within your school, within your college, within your community, to get people to interact. All right. So we, motivated by that logic, decompose economic connectedness for a given person into the sum of three components across the groups where they make friends. So if you think about economic connectedness for a given person, the share of high-income friends they have, we can write it as follows. Take a group g like, say, the high school that you go to, or the church that you're a part of, what fraction of your friends you make there, what share of the members of that group, are high socioeconomic status? How exposed are you to high income people? And then what is your probability of actually befriending them conditional on their presence. So, in particular, we are going to define friending bias just to be clear here is to take the share of high SES friends. You make in a given group tree and divide that by the share of high SES members who are part of Group G. And then we're going to take one minus that. So if you're making friends exactly in proportion to the presence of high SES members, say thirty percent of the people in your group are high SES, and thirty percent of the friends you make are high SES. Then we're going to say friending bias is zero. You're basically sampling at random. If only twenty percent of your friends are SES in that example, then you're under-sampling the I income folks by one-third and friending bias is going to be a third. So if you just work out the math, you can see that if I take the product of these three things in each group and add that up over all the groups in which you participate. I'm going to get to know your overall share of high income friends. Okay. So this decomposition is useful because we can measure each of these components in the Facebook data and then figure out ah, the relative contribution of the exposure channel versus the friending bias channel. Ok, So how do we actually operationalize this in the data? So the key step here to make this work is that we are able to map friendships back to where they are formed. So, for instance, if Markus and I both went to high school, if we were friends on Facebook, and we both went to the same high school within three years of each other. Then we're going to deduce, you know, probably correctly, that we became friends in high school, and in that particular high school that we both reported in the Facebook data.

57:16

So, using that kind of technique, we're able to map the friendships that we see in the Facebook datas to six different types of settings which represent the most common settings where people make friends in the US. Schools, colleges, workplaces, recreational groups, religious groups and neighborhoods. So let me show you now the statistics that we measure in each of these groups, and then show you how important exposure is versus friending bias. So I'm going to start by measuring these friending shares. So here's the first picture there. So here we're looking at the college setting, and we're asking what proportion of your friends do you make in college based on your income level, and what you can see here is, I think, a quite intuitive pattern. High socioeconomic status people tend to make a lot more of their friends in college relative to low socioeconomic status people, partly for the simple reason that they're more likely to go to college to begin with. Now we can look at all of the other settings, workplaces, high schools, and so forth. And you can see, you know what offsets this college pattern where low-income people tend to make their friends. They tend to make their friends at a much higher rate in their own local neighborhood than high-income people do.That, I think you know, maybe it resonates with some of you, it to me is an important finding, because it explains something that we had seen in our earlier work that I never fully understood. Which is that where you grow up seems to matter a lot more. If you're a low-income person than a high-income person, It's a very clear pattern, documented in a number of studies, and this is very consistent with that. I think the social network that you have is much more heavily defined by where you live for lower income folks than higher income folks who tend to be connected with people who are all over the place. Okay, So that gives you a sense of the friending shares by income across these different settings. Now let's turn to the next component that matters which is exposure. So take the set of below median SES people, and ask how exposed are they in their groups to above median SES people in each of these different settings? So this .82 number says, for instance, that on average you have an underrepresentation of about eighteen percent relative to what it would be if you had random exposure in the neighborhood you live in as a low-income person, and that's because we have residential segregation in the US. When people tend to live around low income people. Similarly, for all of these other groups, the one exception is college. If you go to college, you tend to be exposed to a lot of high income people precisely because high-income people tend to be much more likely to go to college to begin with, so that's exposure. And then finally, I come to the third component. I think the novel component that you can measure with this data is friending bias. So this is saying conditional on exposure. How likely are you to befriend somebody from Ah, from a high income background in your group, and in particular, how much do you deviate from that random friending benchmark. So again, starting with this first number here for neighborhoods, this .16 number means sixteen percent friending bias. So if I'm presented with ten high-income people in my group, I'm about sixteen percent less likely to befriend them than I would if I were just making friends at random. To give you some context for that sixteen percent number, remember that overall in the US, when you do the below above median split, there's about a twenty percent underrepresentation of above median friends relative to what you would expect with random friending. So the sixteen percent is quite big in that context, right? If we were to eliminate that friending bias completely, you would go a significant way towards closing the social disconnection.

1:00:55

Raj Chetty: Okay. So in neighborhoods and in a number of these other settings you have quite a substantial amount of friending bias, but there are some real exceptions to that. I think the most important being religious groups in recreational groups where you're seeing very little friending bias, or actually friending bias going the other way, where you're more likely to befriend somebody from a high-income background, and so you know what might be going on there, or some but what we're seeing, the fact is, if you look at people, and who they meet in churches, they're much more likely to befriend somebody from a different socioeconomic class in their church than they are in other settings. It could be because you have something else in common like your faith that could be something else that's different about their setting. What's very clear is that we see the types of friendships people are forming in church are very different from the types of friendships people are forming in other settings, so the setting in which people interact seems to matter quite a bit. So putting together the three components that I just described, friending shares, exposure, and friending bias, we can now use those parameters to quantify the contribution of the key channels that I just described. Exposure versus friending bias in explaining the level of social disconnection in America. So I'm going to show you that result with this chart here, where we start with the level of economic connectedness for low SES people. Remember that number is around .8, about a twenty percent under-representation of high-income friends. For high-income people, that number is about 1.5. They have about fifty percent more high-income friends than the average person. Why do those numbers not average to one? It's because high-income people tend to have more friends to begin with. And so that's why you get this pattern here? Okay. So now I'm going to do the following counterfactual exercise. Let's imagine that low-income people make friends in the same settings as high income people do so rather than making more friends in their neighborhoods. Let's say they make friends at the same rates and colleges as high-income people. But how much of the gap would this close? A very modest amount, because there's a lot of difference in exposure and friending bias across groups within each of these settings, so where you're making friends by itself doesn't explain that much of a gap. Let's now do a second exercise. Let's equate friending rates, and also exposure within each of the groups. So the way to think about this is I'm now going to assume that when you go to college you are exposed to the same number of high income people as high-income kids are currently exposed, and do the same thing for neighborhoods, and so forth. I'm basically going to eliminate segregation on all of the different margins that we're looking at. You can see that this brings the number up to 1.2, closing about half of the gap relative to where high-income people are, and then mechanically, if I were to equate both friending rates, exposure and friending bias, then I would get to where the high income people are. It's just an identity. I'm going to get to that 1.5 number. And so what we learn from this analysis. I think the key takeaway is that about half of the social disconnection between low and high income people in America is driven by a lack of exposure. The fact that low-income people live in different neighborhoods, go to different colleges, go to different schools, go to different churches, et cetera, but then the remaining half is explained by friending bias, lack of interaction, conditional on exposure. So both factors are equally important in response to that poll question initially.

1:04:21

Raj Chetty: And so why is that? I think that's very important from a policy point of view, because it's addressed that even if we were to manage to integrate every school, every college in America, you'd still have half of the social disconnection left between low and high income people.

Markus Brunnermeier: So to put it differently, it's important to organize parties centrally and mix people in the same institution as the spheres there.

Raj Chetty: That that's exactly right, Markus, and that's kind of what I want to transition to and end on. Is it about organizing parties? Or is it about other types of things that we can do to systematically try to address friending bias in the same way that I think we think a lot about how to address exposure. You know, places like Princeton and Harvard devote a lot of effort nowadays to admitting a diverse student body, but conditional on admitting a diverse student body, how do we think about how they are interacting within our institutions? And so in the last few minutes here I want to show that we can think systematically about friending bias in the same way that we often think systematically through potential policy solutions about exposure. So to do that I want to show you that friending bias and exposure vary, not just between these different settings like I've been looking at here, but across groups within settings, and we can learn a lot from that, I think. And I'm going to illustrate that by focusing on the case of high schools. So here is a plot of friending bias and exposure. So the share of high SES students and the degree of friending bias conditional on exposure in each of these schools with one dot per school in America. So again we've released all of these statistics publicly. You can look up your own school in your neighborhood, and what you can see is that there's a lot of variation in schools on both of these dimensions, and importantly, you can take schools that are known to be very diverse. So take, for example, the big public high school in Berkeley, California, near the University of California, Berkeley or Evanston Township High School, which is right next to Northwestern University. These are pretty diverse schools, which have, like a good mix of high and low-income students, roughly fifty fifty. But they also have quite a bit of friending bias. So this is sorted so that higher numbers, higher levels of friending bias are lower on this chart. And so you take a place like at least a town trip, you know, there's quite a substantial amount of friending bias there. Low-income kids are significantly less likely to become friends with high-income kids in their school than they are with the low-income kids in their school, and that is actually a point that has been well documented in prior ethnographic work by sociologists looking at the Evanston High School, which is known to be a very fragmented place along various lines. And so, using this data, we can start to look for patterns. You know, we see that there are some schools that are pretty diverse that have very low-friending bias. There are other schools where you know you have very high levels of exposure like Exeter Academy. You know, aprivate school that predominantly has high-income kids, the friending bias is pretty low there, but there are relatively few low-income kids to begin with. So these two dimensions are both important to think about. In some schools a lack of connectedness might be driven by exposure and other schools, it might be about friending bias.

1:07:31

Now, importantly, friending bias varies for systematic reasons. It's not just like people's preferences vary across places. There are systematic institutional factors that seem to predict the level of friending bias. And so let me give you a few examples here. Here is a plot of levels of friending bias versus the number of students per cohort in the school, the size of the school, and you can see very clearly that friending bias is substantially higher in big schools relative to small schools, and that's a systematic pattern we find in the data. So you know intuitively, Markus, coming back to your party example, if you go to a party with say, five hundred people, you're probably likely to gravitate towards the people who are like you, the people that you know. But if It's a party with ten people, you're probably going to talk to everyone by the end of the evening, and we see that very clearly in the data. Smaller groups tend to foster more cross-interaction, that's something very concrete that could be actionable in terms of policy going forward. Another example, places with more tracking tend to have more cross-class fragmentation, measured here by gifted and talented shares. Another pattern you see is that places with more diversity by income or by race, tend to have more fragmentation across class lines. That actually creates a challenge, because it means that as we integrate schools more into great neighborhoods, we might have endogenously more friending bias arising as well. And so one needs to think about how you achieve that mix of both having diversity and having low levels of friending bias, and I think that can really be done only through careful, deliberate policy efforts. So let me support some.

Markus Brunnermeier: How can sports help in high schools?

Raj Chetty: And so you know, that's another pattern that we see that when we look at these data, might have seen that earlier slide– recreational group groups tend to exhibit much less friending bias. And so things like clubs outside traditional class curriculum, you know, might bring people together across class lines. I think there are other factors like architecture, some people using the data we put out publicly have already shown that things like schools that have uniforms tend to have less friending bias across class lines. So I think there are a number of things here very much like how we've studied exposure in great detail and thought about policies like school integration, college access, college admissions, and so forth, as ways to change exposure. I think we can think systematically about friending bias going forward as well, and I don't think we have the answers yet, but that is what I see as a key area going forward. So let me wrap up in the last two minutes here.

Markus Brunnermeier: Can I just ask one quick question, because Eugene asks, how does this change over time? And let me add to that, if you look at new friends in your dataset, and I'm interested in Covid, how has Covid changed the collectiveness and the social capital? You can probably do this over time as well, and look at new connections.

Raj Chetty: And these are both great questions, not things we've looked at yet. But you know again good examples of what I think we can learn with these data. The long time series, of course, is going to be a little bit difficult to get at, because we didn't have these social network platforms going way far back in time, but certainly over the Covid period. You know, what that gets at, I think more generally markets, which I think is a very interesting question is, how does the shift to interaction online affect these kinds of things relative to or remote kind of interaction versus in-person interaction. How does that change?

Markus Brunnermeier: That was another one of my questions. Is Facebook providing a social good in the sense that, providing more social capital or not?

Raj Chetty: Yes, I know. You know our study here does not speak to that, and you know, I think that remains to be studied. In a sense you can see how there's a lot of potential because you're no longer bound by geography, so you can sort of solve the exposure problem and connect to somebody who's very far away in principle from a very different background. In practice you might expect there's a lot of friending bias that can be generated online because you tend to go into uh groups that have people like you and so forth. So my instinct, you know not empirically, But my intuition is that there's a lot of potential for online interaction to create this type of cross-class connection. But it may not happen, you know, just endogenously. It might be something that has to do with delivery, and so kind of in that vein. Let me wrap up here by talking about how one might actually intervene to increase economic connectedness going forward. And I'm going to skip over some details here. But basically, what we show, in the end of the second paper is that the statistics we release publicly school by schools, zip code by zip code, college by college are sufficiently precise and reliable that you can predict the effects of marginal efforts to say create more integration, using cross-cowork fluctuations and the share of high income peers that someone has how many cross-class connections does that create? It turns out that our measures of funding bias pretty accurately predict whether you're likely to generate more or less cross-class interaction in the schools where we're seeing a lot of friending bias. When you happen to be in a cohort with more high income kids, you become friends with fewer of those kids than in a school with low-friending bias. And so just coming back to the chart that I showed here before you know what that means concretely is, people can take this data, and in a place like Lake Highlands High School in Texas, you know, really think about why is finding bias so high here? And what can we do to tackle that? Whereas in a place like you know, exit, or how do we create more diversity in the student body to admit more low income kids, et cetera. And so, just to show you how schools are already starting to think along these lines, and this is actually something that might be feasible to go forward. We just end with a couple of concrete examples. So here's an example from Berkeley high school where they had figured out before we released this data publicly that the school was incredibly fragmented, that there are essentially multiple divisions within the school of different classes of kids, and they developed an intervention that is very much in line with that small group idea that I had just talked about earlier, where they put kids in smaller houses that cut across different class lines when they enter in their freshman year before tracking them into different groups based on their academic preparation. So we can see going forward whether this kind of thing actually works in creating more cross-class interaction, but it's one illustration of the type of thing one might actually be able to do.

Markus Brunnermeier: There are some high schools where the seating is prescribed. And now, at other high schools, there is free seating in class.

1:14:10

Raj Chetty: Yeah, that's another possibility, you know, even within the classroom thinking about how assignments are made. I think that kind of thing could matter as well. Going outside the school setting, Markus, you know, just to show It's not just about schools, I think. Here's an interesting program that we've recently started to to connect with, which is a gym in Boston that really seeks, I think you know, in our terminology, to directly reduce friending bias. So what they're doing is taking personal trainers who come from disadvantaged backgrounds often with the prior criminal interaction of the criminal justice system. Things like that, And they connect them with clients from affluent backgrounds, CEOS of companies, and so forth. And what they are focused on is not just creating this, providing this personal training service. But if you focus on the text and the bold here, they point out how the people in our program gain access to new networks and opportunities, and they're fundamentally focused on creating this kind of bridging social capital across class lines. And they argue, and again, this could be studied systematically going forward, that this really changes the pathways and the opportunities that people have. So you can think of this as a direct way to tackle a friending bias. So let me wrap up, I hope I've shared a bunch of results. You know there are really two broad messages I'd like you to take away first. I think social capital is measured by economic connectedness. Cross-class interaction in particular appears to be a key media of economic mobility. I've been studying these issues for a number of years, and this turns out to be the factor that's the strongest predictor of economic mobility that we or others have found debate. So I think there's really something we're focusing on going forward there and then second, I think if we think about how to shape economic connectedness. It's not just about segregation, but it's also about friending bias, and both of those can be shaped by policy. Zooming out a bit from this particular data, and this study more generally, my sense is that social connections appear central in many recent programs that have shown promise and increasing upward mobility. So in other work we've done on our team and helping families move to higher opportunity neighborhoods or other work people are doing on drop training programs, all totally different spheres, the policies that look like they're successful are ones that don't just give people resources like: Here's a housing voucher. Or here's a set of technical skills, good luck, and finding a job or finding a better place to live. But in addition they provide some social support. They provide social capital if you will. That helps you make use of those resources. And so, I think, going forward, at a very big picture level, I think, as we think about how to design policies, to increase upward mobility and reduce inequality in society, traditionally, in economics we tend to focus a lot on resources and incentives and things like that, and I think those things matter a lot. But I think what we're seeing from this recent body of work is that it's equally important to focus on socioeconomic connections and social capital. And my hope is by understanding that better, partly, perhaps, using the data that we've released publicly here, which we hope people will find useful. You can download it at socialcapital.org, will be able to target such interventions and study their impacts more precisely, going forward. So thanks so much.

1:17:26

Markus Brunnermeier: Thanks a lot Raj. This was fascinating. We learned a lot about social capital and social mobility, and I hope that many of you will download the data and play with the graphs, but also work on the data and develop this whole literature further, I think that's a great service you're doing as well to provide the data to everybody, to analyze them further. I have many new additional questions, but we are already running a little bit over time, so let's leave it at that. I hope I see you soon again, and thanks again for doing this, and I will probably see you tomorrow at Princeton when you come tomorrow.

Raj Chetty: Thanks so much for it. My pleasure,

Markus Brunnermeier: Bye, bye, everybody, and I hope to see you soon again for the next webinar. Thanks for being with us.