

## Probable Causation, Episode 25: Sara Heller

**Jennifer** [00:00:07] Hello and welcome to Probable Causation, a show about law, economics and crime. I'm your host, Jennifer Doleac of Texas A&M University, where I'm an Economics Professor and the Director of the Justice Tech Lab.

**Jennifer** [00:00:17] My guest this week is Sara Heller. Sara is an Assistant Professor of Economics at the University of Michigan. Sara, welcome to the show.

**Sara** [00:00:24] Thanks so much for having me.

**Jennifer** [00:00:26] We are going to talk today about your research on summer jobs for teens as a violence prevention strategy. But to kick things off, could you tell us about your research expertise and how you became interested in this topic?

**Sara** [00:00:38] Sure. So broadly, I'm interested in ways that policy can help improve outcomes for disadvantaged youth. And so that means understanding how adolescents and young adults make decisions about school, about crime, about work, and how their thinking interacts with their environments and their social context. So in terms of the sort of type of program we'll talk about today, which is summer jobs programs, I first got interested in it in graduate school when I was reading a bunch of research about active labor market programs, so programs that offer job search assistance, job placement, job training. And we know a lot about those types of programs, but I realized through reading that research that we actually knew very little about summer youth employment programs, even though they've been operating all over the country, we're pretty widespread for basically the last 50 years. And so when you're a graduate student and you realize there's something that's kind of a big deal, we've been spending a lot of money on it, and we know very little about it, that's sort of an exciting opportunity to to think about doing some research.

**Jennifer** [00:01:37] Indeed. So the main paper we're going to talk about is titled "Summer Jobs Reduce Violence Among Disadvantaged Youth," and it was published in Science in 2014. You've continued to study the effects of summer jobs and more recent work, and we'll talk more about that later. But let's start with the original experiment. So you randomly assigned summer jobs to teens in Chicago. Tell us about that summer jobs program. What did the program look like, and who did it target?

**Sara** [00:02:01] Sure. So Chicago offers a huge range of summer programming that ranges in what it looks like, and the piece we evaluated is a piece that we sort of carved off in order to evaluate with a focus on violence prevention. So that means it was offered to youth at 13 high violence, high schools, Chicago public schools. And in a lot of ways, it's very similar to other types of summer jobs programs across the country. So youth worked five hours a day, five days a week for a minimum wage. In this first study, they were all nonprofit and government jobs. So things like summer camp counselors, working in an alderman's office, and so forth.

**Sara** [00:02:39] But there were a couple elements of the program that the government agency that ran the program decided to add because of the population they were targeting and the sort of focus on violence prevention. One of those pieces was an adult mentor. So this is an adult who works with about 10 youth for 1 adult whose job it was to help them deal with their barriers to employment. So whether that was figuring out transportation to work, helping them get the paperwork together they needed or the identification they didn't

have in order to work or helping them navigate conflicts with supervisors, they were basically there to help make sure that youth succeeded. And then some of the youth also participated in a social emotional learning curriculum. So that was sort of a piece of the program based around cognitive behavioral therapy principles, helping with social information processing, emotional regulation, and that kind of thing.

**Jennifer** [00:03:32] So you mentioned there was room for more research here, but before the study, what had we known about the effectiveness of jobs programs like this?

**Sara** [00:03:39] Surprisingly little given how widespread they were. So we know a lot about active labor market programs in general. So these kinds of job search assistance, job placement, job training programs, they're actually one of the most evaluated social programs out there. We know that they improve employment for adult women. They're sort of mixed evidence on whether they help improve employment for adult men. The evidence on youth is even more mixed. But when we think about the summer jobs program, so that are targeted at youth who generally are still in high school, that's why it's offered during the summer, we really only had basically observational evidence on summer jobs programs.

**Sara** [00:04:17] So studies that compared participants to non-participants. There was a one program back in the 80s that was evaluated as an experiment, but as you sort of dig into it, you figure out that they offered both the treatment and the control groups jobs and training. The only difference between the two groups was a life skills curriculum. So that's not actually evaluating the effects of offering a job and some job training.

**Sara** [00:04:40] So going into this - starting this study, we basically only had that observational evidence which was generally promising, but not entirely convincing, given that the people, the kinds of people who show up for these programs might be different from the kinds of people who don't show up. So right before my paper came out, there was an experiment that came out about New York City's summer youth employment programs, talking about the academic effects of New York's program. And so there Jacob Leos-Urbel found sort of small increases in school outcomes on the order of sort of one or two extra days of attendance for participants in the program. And I think there was a bunch of other work going on concurrently. But the - at the time of the study, that's basically what we knew, some promising observational evidence and some experimental evidence on schooling, but nothing about crime.

**Jennifer** [00:05:35] So to measure the causal effect of this program in Chicago, you ran a randomized experiment, something we wish we could do in many contexts, but isn't always possible. And as many of our listeners know, convincing practitioners to randomize access to their services isn't easy, even in cases when doing so is logistically feasible. So how did this experiment come about?

**Sara** [00:05:55] It's kind of an interesting story. So very early in graduate school, I had met with a couple of the people who run the program in the Daley administration. So this was back when Mayor Daley was mayor of Chicago. I told them I was interested in this topic. I was interested in figuring out if there are opportunities for evaluation. They said, oh, yeah, we might be interested. Thanks very much. And I sort of went on my merry way.

**Sara** [00:06:18] And what it took to start to get things off the ground was Rahm Emanuel getting elected. And so he came to office, and he gave all of his staff sort of very explicit directive that he actually wanted to know if the stuff that he was spending money on, what worked or didn't work. And so at that point, the staff sort of remembered that I had talked

to them about figuring out what works and what doesn't work. And they called me back in to discuss the possibility of an evaluation. And I think the thing that really made the study happen at the end of the day was the leadership of Evelyn Diaz, who was the Commissioner of the Department of Family and Support Services at the time, which is the agency that runs the program. She really made it all possible.

**Sara** [00:07:04] So she was someone who cared a lot about evidence, who understood the benefits of randomizing a study, and who also understood that in a world of social services where you basically never have enough money to generate enough program slots to serve everyone you think could possibly benefit from the program, you always have to make choices about allocating slots. The limited number of slots are getting allocated somehow, and a lottery is potentially a fairer way to allocate them than however they would get allocated otherwise, be that first come, first serve or it's the people who sort of show up first who get it or nepotism because it's sort of who you know. And so she was willing to make the tough choices, to make it work, and really in general was just an absolutely extraordinary partner who helped get the project off the ground and made it a success at the end of the day.

**Jennifer** [00:07:55] Okay, so you had awesome policy partners that allowed you to randomly assigned kids who want to participate in this program into 3 groups. So you have a control group, where kids don't get a job, but then you have 2 different treatment groups. So tell us about each of those groups and why you did this.

**Sara** [00:08:10] So the difference between the 2 groups is the social-emotional learning curriculum piece that I told you about earlier. So one of the groups worked for 5 hours a day, 5 days a week, and that was it. The other group worked 2 fewer hours a week - sorry, the other group worked 2 fewer hours a day, and instead replaced that with this social-emotional learning curriculum that's based around cognitive behavioral therapy principles, so CBT is what we sort of call that for short.

**Sara** [00:08:42] And the reason we did that is that there had been a lot of sort of mixed evidence about active labor market programs for youth, and a lot of sort of failed attempts to offer programing that would improve employment in the future that didn't really seem to pan out. And so there was a hypothesis floating around about there that part of the reason that disconnected youth were not benefiting from these programs is that they lack the soft skills to benefit from the programs, whether that was about emotional regulation or self-control or communication or just the other types of skills that might really let you engage with programing, benefit from it, grow from it, have it set you on a better employment trajectory, then maybe they needed a little bit of training in those skills. And so our thought was to test whether adding something like that, adding a CBT based curriculum, could provide some of those soft skills in a way that would help the youth get more out of the program than they would otherwise. So we wanted to sort of see if that would help expand the program impact.

**Jennifer** [00:09:44] Okay, so randomizing kids across groups allows you to measure the causal effect of the program. That's one piece of the puzzle. But you also needed high quality data to find out what happened to the kids in each of these groups. You were able to get a bunch of administrative data to measure the effectiveness of this program. So tell us about those data sets and the outcome measures you're interested in here.

**Sara** [00:10:04] So this is one of the benefits of working in Chicago is they have extraordinary administrative data access. It was not all that easy to get. So I spent

basically my entire fourth year of graduate school arguing with the lawyers to figure out exactly how we could protect everyone's privacy, but still learn what we needed to learn. And at the end of the day, we got a lot of data.

**Sara** [00:10:25] So the main outcome measures we focused on were individual level arrest records from the Chicago Police Department. So we can measure each time a youth gets arrested and importantly, why they were arrested. So I sort the arrest by crime type to think about violent crime, property crime, drug crime, and other types of crime. Part of the reason I do that is that there's a lot of evidence from other studies that different types of crime sometimes move in different directions, probably because there's different underlying causes. And so if you change one cause, but not the other, things might move in opposite ways. And I didn't want to aggregate that all together. I wanted to be able to sort those pieces out. So the Chicago Police Department arrest records are the main outcome I'm interested in. I also match youth to their Chicago public school record so I can see attendance, enrollment, GPA, other school outcomes. I, of course, have program records from the program about participation, and I know a little bit about the neighborhood youth live in from the American Community Survey.

**Jennifer** [00:11:28] Let's talk about the main results. What do you find is the effect of being assigned to the treatment groups on kids' academic outcomes and criminal behavior?

**Sara** [00:11:36] So in the 16 months after random assignment, I find that there are 4 fewer violent crime arrests per 100 youth offered the program. And to give you a sense of that magnitude, that's a 43 percent decline relative to the control group. So that's a really substantively large decline in violent crime arrests. I don't find any change in other types of crime or in any measures of school attendance or performance. So it looks like it's really the violent crime effect that is the the main change due to the program.

**Jennifer** [00:12:11] And that's nice because violent crime is very expensive. So that's a big effect on an important outcome.

**Sara** [00:12:16] And it was the focus of the policymakers running the program as well. That's the thing that they really wanted to target.

**Jennifer** [00:12:22] Yeah. And you don't find any difference between the 2 treatment groups. So that is receiving CBT on top of the job did not have any additional effects. I imagine this is a bit surprising to you. Why do you think there was no difference?

**Sara** [00:12:35] So I think there's sort of a boring statistical reason and then maybe a more substantive hypothesis. The statistical reason is that, of course, as - rather than comparing the entire treatment group to the control group, when you look at one treatment group and the other treatment group versus the control group, you're sort of slicing your treatment sample in half. And smaller sample size does make it harder to find and sort of detect the differences. And so part of the answer might just be because the sample was not big enough to really tell. That said, the point estimates or the actual sort of estimate of the effects are pretty similar across the two groups. So even though there are big standard errors, so there's some uncertainty. All the indications are that the groups look pretty similar. And in fact, the violent decline is statistically significant in both groups separately. So it's not like the sample was so small that we couldn't see the violence decline in both groups, which means certainly the CBT piece wasn't necessary to cause the violence

decline. It happened in both places, even when youth were not participating in that curriculum.

**Sara** [00:13:37] And so, as you suggest, that is kind of surprising because we have this other evidence that suggests these programs by themselves can reduce violence. And so part of what we learn from this, I think, is that if the CBT is not necessary, it must be something that is happening in both of the groups that's driving the effect. And I think one important piece of that is to realize it's not like it's the exact same program, plus the CBT, that the youth in the CBT group are actually working fewer hours. They're replacing two hours of work with the CBT. And so you might think that maybe the two intervention strategies, the work and the CBT are interchangeable, that they're actually doing something similar.

**Sara** [00:14:21] And one of the reasons that I think this is potentially plausible is back when the program was happening, I went on some site visits. I talked to the people in the program and some of the employers. And in talking to one of the employers, he told me this story and he said, look, the biggest problem these kids have when they come, when they first show up, is that they can't take constructive criticism. That you tell them, hey, you need to wear closed-toed shoes to work, and they blow up at you. They're in your face yelling at you. And so he said, my job as an employer is to teach them not to do that, to teach them how to take that constructive criticism. And if you think about that, that sounds a lot like what CBT is trying to do too. To help people sort of slow down, catch their hot reactions, think a little bit about their own thinking and about what the person is trying to communicate to them, and manage that sort of thinking and that emotion in a more constructive way. And so it might be that the kids - it's not necessarily the case that the CBT lessons don't matter. It's just that they might be learning those same lessons from the kinds of employers who choose to participate in these programs because the employers themselves are also interested in youth development.

**Jennifer** [00:15:32] That's really interesting. What else were you able to say about the mechanisms driving these large effects based on the data from this first study?

**Sara** [00:15:40] So frustratingly, it's a little bit easier for us to rule out mechanisms than it is to rule them in and to say for sure what's going on. But there are a couple of things that we can rule out. So one, like I told you, is that there weren't any changes in school performance. So it's not the case that youth are getting much more engaged in school either because they sort of learn about the value of school in the labor market or because they're more motivated or are feeling sort of better or self-efficacy. That's not what's going on.

**Sara** [00:16:09] The other thing we can rule out is - one thing you might have thought is, well, of course, kids are getting arrested less during the summer because they're busier. They just have less time during the summer. And so we would sort of call that an incapacitation effect in that they're incapacitated while they're at work. And while you're physically working, you can't also be physically getting yourself into trouble. And so if that were the case, we would expect to see the arrest drop focused on the summer months. Right. So you might see a big drop during the summer months and then everyone would go back to the behavior as usual after the summer months. That's not actually what we see. So there is a sort of proportionately large but not statistically significant decline during the summer months, partly because it's only three months and just not a lot of crime happens in any given three months. But the important part is that we then see that crime decline, the violent crime decline, continuing after the program is over. And in fact, if you

throw out the summer months entirely, you still have a significant decline in violent crime. So the kids are taking something with them that changes their behavior into the future. It's still a little bit of an open question about what that something is.

**Jennifer** [00:17:20] So to dig into this further, you continued the study with a second cohort from Chicago as well as some additional data. So tell us about that second cohort. How did it differ from the first?

**Sara** [00:17:31] So Commissioner Diaz, who was running the program, had seen some of our very, very early hints of the violent decline. So you have to sort of think about this in calendar time is that the summer of 2012, we implemented the program. We matched the data and did some very early analysis in those months over the fall of 2012. And starting in December, January of 2013, the city has to start making decisions about what to do next summer. Now, as anyone who has - is either is an academic or has worked with academic knows, academics don't work on that timeline. Right. It's not like, oh, three months later, we know the answer. Everything's fine. We need to let some time accrue to actually sort of let the outcomes realize and then analyze them and then make sure our analysis is right. But we had to make a decision about what to do the following summer. And so what Commissioner Diaz said is well, if there's even a possibility that the program might actually decrease violence, I want to focus on a population who is at higher risk of violence than the one we served. I really want to think about whether this is going to work for the more disconnected population.

**Sara** [00:18:39] And I told her at the time that I didn't think it would. And that's partly because of the other evidence on employment programs for more disconnected youth. It's partly because I sort of had a hypothesis that if the first program actually was decreasing violence, maybe it had to do with the fact that prevention is easier than remediation, that getting to youth early before they face unemployment, before they become disconnected from the labor market might be an easier time to intervene than after they've experienced some unemployment or been out of the labor market and out of school. And so I told her, I don't think that'll work, but it's an empirical question and we can test it. And so the second cohort we recruited - we changed the eligibility criteria to recruit a somewhat more disconnected population. So first, we focused all on boys in that second cohort because they're disproportionately involved in violence. And we recruited about half the sample directly from criminal justice agencies. So juvenile and adult probation, the Juvenile Detention Center, a nonprofit that was working with people, with kids who were involved in the criminal justice system. It was never a condition of their probation. It was never required. We just offered the opportunity to apply to - about half the sample from that sort of pathway. And the other half of the sample was more like our original sample. And so now we had this variation in observable characteristics that sets us up to explore treatment heterogeneity or to explore for whom the program works. And so that was the focus of the question of the second cohort, is figuring out who does the program really work for and who does it not work for?

**Jennifer** [00:20:30] And then what additional outcome data did you gather for the subsequent analysis?

**Sara** [00:20:34] So two things changed. One is that we got access to Illinois state police records instead of the Chicago Police Department records. So we're measuring a sort of broader set of arrests and catching any arrest that might happen outside the city of Chicago. And the second piece is that we added employment data. So we got access to unemployment insurance records, which are kept by a state agency. And so that's the

agency that when unemployment insurance taxes get withheld from your paycheck, that's who keeps track of that, so they can see and measure formal work quarterly across the year. So it's not a perfect measure of employment. It's missing informal work, which ethnographic studies suggest might be a big deal for this population, it's missing the self-employments, but it's likely to be more complete than anything we would be able to get through surveys, which would be large and expensive and probably have a lot of missing data.

**Jennifer** [00:21:28] And formal employment matters for a variety of reasons. Right. We like these kids to be involved in that type of work all else equal.

**Sara** [00:21:37] That's right. And I think when when policymakers think about improving future employment outcomes, they're thinking about the sort of formal employment sector.

**Jennifer** [00:21:43] Yeah. Okay, so all of this became a second paper titled "Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs." It's coauthored with Jonathan Davis and is forthcoming at the Review of Economics and Statistics. So first, what did you find when you added the second cohort and those employment data? Were the results in line with the first experiment?

**Sara** [00:22:04] So recall that I told Commissioner Diaz that I didn't think the program would work for that second cohort. I was totally wrong. So we actually basically replicated the violence drop that we saw in the first cohort. So in that first cohort, it was - if you look just at the first year, it was a forty two percent decline, which was 4 fewer arrests per 100 youth offered the program. The second cohort, so the more disconnected one, was a 33% decline in violent crime arrests. So basically in the same neighborhood as that 42%. Now as a bigger number, in terms of the drop, so there were actually 8 fewer arrests per 100 youth offered the program because the youth were more criminally active, but proportionally it was basically the same drop. And so that by itself is an important finding. I think we tend to undervalue replication in economics. But the fact that we went out, did this again with a different population, and found basically the same thing, I find extremely heartening, in terms of it not just being a sort of fluke case. And then publication bias means you can publish your first big estimate and then you try to get it again and you can't. So I think the replication by itself is important. We could now follow youth for a little bit longer time, and so we could see that the drop is basically concentrated in that first year. If you look out the second year or for some of them, we could - the first cohort we could see for the third year. We don't see that the drop continues accruing.

**Sara** [00:23:34] Now it doesn't catch back up. So the cumulative effect is still a decline in the sort of total number of violent crime arrests. But it's not an effect that sort of keeps accruing forever. So that's on the crime front.

**Sara** [00:23:48] On the employment front, we found no average change in employment. So the program was not doing what I think a lot of policymakers tend to think about summer jobs programs doing, which is preparing youth for the workforce, setting them on a better employment trajectory, teaching them the sort of introductory work skills that they need to start off their entry into the labor force more strongly. At least in the two to three years that we can follow youth, there were no average changes in employment.

**Jennifer** [00:24:16] That raises questions about what the potential mechanisms are here. Right. So you and Jonathan actually use machine learning to try to say more about the heterogeneity in the effects itself, in the hopes of shedding light on potential mechanisms

since it on its face seems like employment is not what's driving this, which I think is probably a surprise to a lot of listeners. So what was the motivation for this machine learning approach? And give us a brief summary of what it involved.

**Sara** [00:24:42] So the motivation is this sort of focus on the question of for whom does the program work? We wanted to answer that question. And, you know, you might think even though there's zero average employment effects, there could be some subgroup that's really benefiting in terms of employments who could be driving the decline in violent crime. So it might be that it's the sort of work driving the violence decline, but only for a subgroup, so that when you combine everyone's data together, the employment effect is not big enough to sort of show up in our analyses. So we wanted to think about who's benefiting and who's not on these different outcomes. Now, typically, when we do these kinds of heterogeneity analyses in economics or in other disciplines even, we sort of think about testing two subgroups at a time. So we might estimate the effect for boys and for girls and test if they're different or for people above and below the federal poverty line and test if they're different and so on across all of the things that we can measure one group at a time.

[00:25:42] Now, there's two potential problems with that strategy. One is that as anyone who has sort of taken an introductory statistics and hypothesis testing class knows, more tests means a higher probability of chance results. Right. Sometimes you're going to find statistically significant results just because you've run so many tests. And the more tests you run, the more likely that is. And then you might end up trying to interpret these subgroup differences that are really just an artifact of the fact that you've run so many tests. So that's one reason we might not like the approach of running a bunch of these tests over and over across different characteristics. The other reason is that the real world might be more complicated than these sort of one way tests boys versus girls, poor versus not poor, that imagine it's the case that the people who really benefit from the program are boys between the ages of 14 and 16 who live in areas where the median neighborhood income is below \$33,000. If there's sort of a combination of characteristics that define who's benefiting, we are never going to find that in our standard way of testing. In fact, if - for the people who sort of do these things, if you think about running that like three or four way interaction and presenting that to a roomful of economists, they would sort of laugh you out of the room about how much data mining you've done and you've just sort of dug around until you've found enough stars. And so we didn't want to do that. We wanted to take a sort of different approach to estimating treatment heterogeneity that takes advantage of some of the new advances in machine learning.

**Sara** [00:27:16] And so we used a newly developed machine learning algorithm and this was not ours. This is thanks to some very clever researchers. And the algorithm basically searches in the data for responsive subgroups. And so it does it in a principled way. So it's holding out part of the sample to avoid the sort of overfitting problems that we might worry about trying to avoid capitalizing on the chance variation in your data. And it allows for very flexible combinations of the characteristics we can observe. So at least in theory, the algorithm could find that group of 14 to 16 year olds living in particular areas who are of a particular gender who are benefiting. So it sort of addresses both of those problems. And so that was our basic approach to let the algorithm loose and search for treatment heterogeneity across all of the things we could observe about the kids.

**Jennifer** [00:28:11] And so what did you find? Which groups were most affected by the summer jobs program?



**Sara** [00:28:15] So we found that there was not any detectable heterogeneity on violence. So basically everyone in the sample seems to respond in terms of violent crime arrest. Everyone goes down. Now, that might be because our sample is not big enough. These machine learning methods really are designed for big data. So maybe our data aren't quite big enough. But regardless, we basically saw everyone's benefiting on violence. The place where we did pick up some heterogeneity is in employment. So we found a subgroup where some people benefit in terms of employment and in fact, by quite a lot. So we found a 15 percentage point increase in employment rates, which is about a 44% increase over the control group.

**Sara** [00:28:59] So some people are benefiting a lot in terms of employment and some not at all. The people who are benefiting. If you look across the characteristics that describe them, they're younger, they're more attached to school, they tend to be more Hispanic and more female. And so if you think about especially the younger and more attached to school, that's sort of the opposite from the population that we tend to think about serving with employment programs. So most of the employment programs we've been running are serving that out of school, out of work youth. Right. People who are disconnected. And in fact, it's the opposite type of person who, at least in our sample, seems to be benefiting on employment, the people who are still in school and a little bit better off on various characteristics.

**Jennifer** [00:29:39] Yes, it kind of goes back to what you were saying earlier about prevention versus rehabilitation. Prevention is easier in many cases. So what does that heterogeneity tell us about how the program is working?

**Sara** [00:29:54] Yes. So we set this up sort of saying, well, maybe if there's a group that is improving on employment, they're the ones driving the violent crime. Once again, this is a sort of hypothesis that we were excited about going in and it doesn't find a lot of support in the data. So, in fact, if you look at that group of employment benefitters, so we have this machine learning prediction of how much someone is going to benefit on employment, so we can look at the sort of benefitters and the non-benefitters. If you look at the people whose employment is improving, the violence decline is not concentrated among them. It's declining even among the group who is not benefiting in terms of employment, which means it doesn't seem like it's the employment that's really driving the violence decline. And if anything, what we see among the employment benefitters is a little bit of longer term increase in property crime arrests. And so that might be more consistent with an idea of, well, if you're working more, you might be traveling more to wealthier neighborhoods. You're around more opportunities for theft. But it's - it certainly doesn't seem consistent with the idea that employment is causing the the violence decline.

**Jennifer** [00:30:57] So Chicago isn't the only place that experiments with summer jobs programs, you mentioned New York. I know that you've also continued to work in Chicago and there was a Boston study, too. So what have those other studies found?

**Sara** [00:31:10] This is, I think, one of the most exciting things about this topic, because oftentimes when you test a program or an intervention type in different settings, you get very different results and then have to think very carefully about what's different about the setting, what's different about the population, what's different about the specific program implementation details. Here across these different settings, everyone is finding a very consistent story in general. So the New York - there have been a couple of different New York studies finding declines in incarceration, even declines in mortality that they, it looks like is driven by homicide. So that's very consistent with the idea that involvement in

violence is going down. Boston found a decline in violence. There - they also find a decline in property crime. And so I think we're seeing really consistent evidence that the program type, regardless of the sort of specific details, is reducing violence and that it doesn't seem to be improving employment on average. Right. So the different studies that have measured employment don't find a lot of improvement in employment, at least on average.

**Sara** [00:32:19] I think the evidence is a little bit less consistent in terms of the schooling outcomes. So the first New York study and the Boston study both find some improvements in schooling outcomes. Another New York study and all of the Chicago stuff has found no improvement in schooling. So it's a little bit less consistent there. There we might need to sort of start parsing the details a little bit. But in general, I would say, and including some other non-experimental studies that have happened. So there's a non experimental study in Detroit that found some positive schooling impacts as well. So I think in general, what I take away from this literature is that unlike in some other settings, the details may not be super crucial, that it seems like these programs consistently reduce violence.

**Jennifer** [00:33:09] Do you have any other thoughts on why the programs don't seem to be having big impacts on employment? It feels like giving kids experience in a summer job like that feels like the easiest margin to move in many ways. Is it something about them working in the informal sector or there just aren't jobs available or do you have any any ideas about that?

**Sara** [00:33:31] So I think it is a little bit of a puzzle. Right. Because we know that early work experience matters in general for employment trajectories and this is early work experience and we can measure employment in the control group and it's much lower. So it seems like it should matter. I think the thing you mentioned about the sector people are working in is a potential hypothesis. Then one of the New York studies did a sort of non-experimental comparison within the experiment that suggests that the youth who are working in the jobs that look more like private sector jobs are maybe doing less badly, although not necessarily improving, than people who are working in the nonprofit sector. So that's a possibility. I think it's possible that depending on the program in the city, there could be some stigma associated with these programs. So if youth aren't listing them on their resume, as I worked in an alderman's office, but are instead saying I worked as part of this program, employers might have sort of some biases asks about the kinds of youth who are participating in these programs.

**Sara** [00:34:40] One of the authors of the New York study, Judd Kessler and I, are working on a study to test a different hypothesis, which is sort of about - pushing on that information in the labor market. Right. It might be that employers don't know how to interpret this as whether it's real job experience or not, or youth are not listing it in the right way. There might be some sort of information frictions that are keeping youth from benefiting from their job experience. And so we're testing the effects of adding a letter of recommendation. So we surveyed a bunch of the youths' supervisors to get personalized feedback about the youth. We turn those into a letter of recommendation and we're going to see if those - that additional piece of information helps the youth benefit in the labor market. But we don't have the answer yet.

**Jennifer** [00:35:25] Oh, that's great. And then in terms of the heterogeneity, have the other studies been able to look at heterogeneity enough to see if the findings are consistent with what you're finding in Chicago, that it's like the less disattached kids that are benefiting most?

**Sara** [00:35:40] So most of the other studies are sort of serving the more classic program populations who are still engaged in high school, so they have a little bit less variation in those characteristics because they didn't set out to test that. So it's a hard question to answer. I think if you look across the patterns of people doing those sort of one way interaction tests, it's relatively consistent. But I think it's - it would be sort of an area of important future research to sort of take what we're finding and maybe the one way interactions other people are finding, and to set out with the intention of testing that and recruiting your program population accordingly, so that you have enough variation in the characteristics to be able to answer that question.

**Jennifer** [00:36:26] So putting it all together, the results of your work in Chicago and the other studies we talked about, what are the policy implications of all this?

**Sara** [00:36:34] I think if you're a policymaker interested in reducing violence among young people, these programs seem like a good option. That it doesn't seem to matter - right, so the New York program doesn't have the mentoring piece. The - not everyone has a CBT piece in the same way. The - you know, the labor markets are different. The youth populations are a little bit different. And yet everywhere it's been rigorously tested, the program seems to be reducing violent crime. That is by itself, I think, a really important policy implication.

**Sara** [00:37:06] I think it's important to realize the program is not a panacea. Right. Even the treatment group is still offending some. It's not like it's getting rid of all violent crime and the effect doesn't keep accruing forever. So it's not a permanent solution for everybody, but it does seem like a cost effective choice. And this is something that I think is often underappreciated in conversations about fadeout. Right. So there's a lot of social programs where the effects don't persist forever with some outcomes that might be really disappointing. Right. So if we're trying to increase test scores and we do it in grade three, but then by grade four the control group, looks just like the treatment group, we can think about that as real fadeout. In the case of violence, if you're preventing violent crime or violent crime arrests, even for just a year, you've still prevented that level of violence. It's not like if then in the second year the rates of offending go back to what they were before, you have - you've no longer prevented that violence, right? You've still stopped violence for a period of time, which is an incredibly socially costly outcome.

**Sara** [00:38:13] So, like I said, I think it's a potentially good policy option for policymakers thinking about that. But it's also important to think about targeting. Right. So even though we found violence declines for basically every subgroup we could look at, that's partly because of the population. At least in Chicago we were working with mostly living on the south and west side of Chicago, where there is a lot of risk for violent crime. I think sometimes policymakers might sort of see program effects like this. And, you know, it kind of gets the stamp. This is an evidence based program. And so we should do it everywhere for everyone all the time. I think that would be a mistake, right. That if the main effect of these programs is to reduce violence, you certainly need to be serving youth who are at risk of engaging in violent behavior. And so thinking carefully about the targeting could be important, which is not to say I think these programs should serve exclusively people who are at very high risk of engaging in violent crime, because I think we don't understand enough yet about the peer effects. Right. It could genuinely change the effect of the program if you really significantly change the composition of the program.

**Jennifer** [00:39:23] Just one more thought on the fadeout. So I think you say in the paper that one possible explanation is that, you know, kids just age out of violent crime. And so it

might just be something where, like you're seeing the reduction and then by the time you're catching up, it's because basically violence for everyone is reducing. Is that consistent with the numbers you're seeing?

**Sara** [00:39:40] That's right. So the rates of violence, even in the control group are sort of, I think, dropping by about 1/2 in the later years. So it is certainly true that as youth age, they are aging out of violent crime and so you just get less violent crime anyway. And so there might just sort of be less room for reduction there.

**Jennifer** [00:39:58] Yeah. And I - so violent crime is expensive, mortality is also expensive. You mentioned the New York's study finding a reduction in mortality. Do you happen to have ballpark numbers for what these programs cost? I imagine once you reduce those types of outcomes, it's not - it doesn't take much to be cost effective, but how cost effective are summer jobs programs?

**Sara** [00:40:18] So it's very hard to sort of put a price, especially on the loss of life and on violent crime. And the program costs vary a little bit place to place. So Chicago's was a little bit more expensive, partly because the mentors were paid. And so that's a more expensive part of the program. There the administrative cost of the program was about \$3,000. Now, if you think about some of those costs are wages, and so that's a cost to the government, it shows up in the government budget, but it's also a transfer to the youth. And then the youth and their family are benefiting from those wages. So oftentimes economists will think about the wages, not necessarily as a net social cost, but as a transfer, which then leaves a sort of net social cost of the program of somewhere, you know, \$1,400-1,600 in there. I think the Boston and New York programs tend to be a little bit cheaper. I don't remember the exact numbers, somewhere between, I think \$1,500-\$2,000. And so I think basically all of the efforts to put the dollar value on violent crime suggests that the program is most likely generating benefits that justify the cost. And potentially, if you sort of include the mortality piece, generating benefits that are way, way, way larger than the costs.

**Jennifer** [00:41:38] So what's the research frontier here? I know you're continuing to work in this space. So what are the big open questions that remain that you and others will be thinking about in the years ahead?

**Sara** [00:41:48] I mean, for me, the biggest question is why? Why are these programs reducing violence so consistently? I think a lot of the hypothesis we've sort of tried to test turn out to not be the answer. And so thinking about what's left I think seems important. And people doing - are doing that in a variety of ways. I'm trying to test the role of different pieces of the programs or test a version with and without the mentor. In Philadelphia, I have a study that's going to be able to look at related outcomes like measures of mental health and family stress in terms of reports of child abuse and neglect. Thinking about, you know, is the money sort of helping to reduce family stress? Is it reducing depression and anxiety and so forth?

**Sara** [00:42:32] In Boston, Alicia Modestino ran some surveys to try to get at sort of why and what's going on. I think a range of other people are sort of trying to work on varying different pieces of the program, thinking about the role of job type like we talked about. So I think - if you're actually operating these programs, you want to know, does it matter what the kids do? Do I have to find particular types of jobs for them? I've tried to get a lot of jurisdictions to be able to randomize jobs, but everyone says no because they say, and I think probably correctly, that a big piece of the value added of the program operators is

doing that match between kids and jobs. And they don't want to sort of give that up. But I think questions about the role of job type and sector and does it matter what the kids do are important. And I think whether peers matter is important. So if you think about targeting youth who are at higher risk of violence, can you do that exclusively or is that going to change the effect of the program? Because part of what's going on is interaction with a diverse set of peers.

**Jennifer** [00:43:36] My guest today has been Sara Heller from the University of Michigan. Sara, thanks so much for talking with me.

**Sara** [00:43:41] Thank you very much.

**Jennifer** [00:43:48] You can find links to all the research we discussed today on our website, [probablecausation.com](http://probablecausation.com). You can also subscribe to the show there or wherever you get your podcasts to make sure you don't miss a single episode. Big thanks to Emergent Ventures for supporting the show and thanks also to our Patreon subscribers. This show is listener supported. So if you enjoy the podcast, then please consider contributing via Patreon. You can find a link on our website. Our sound engineer is Caroline Hockenbury with production assistance from Elizabeth Pancotti. Our music is by Werner and our logo is designed by Carrie Throckmorton. Thanks for listening and I'll talk to you in two weeks.