

# Consumer Finance Monitor (Season 4, Episode 46): CFPB Director Chopra's Targeting of "Digital" or "Algorithmic" Redlining: What It Means for Providers of Technology and Consumer Credit

Speakers: Alan Kaplinsky and Chris Willis

Alan Kaplinsky:

Welcome to Consumer Finance Monitor podcast, where we explore important new developments in the world of Consumer Financial Services. I'm Alan Kaplinsky, the former practice group leader of the Consumer Financial Services group at Ballard Spahr, and I will be hosting our show this morning. I'm very pleased to have as my guest this morning, it's a little awkward calling him a guest, because he's really a colleague. My good friend and colleague, Chris Willis.

Alan Kaplinsky:

Chris is a current co-chair of the practice group, the Consumer Financial Services group. And also, among the many hats that he wears, he is very much involved in issues pertaining to fair lending, fair banking and redlining. And redlining is the topic du jour. So I first, Chris, want to welcome you back. I know we have talked previously quite a bit about regular redlining, if I could refer to it as that. Well, I'm sure I'll get an explanation from you during the course of our interview about the difference between regular redlining and today's topic, digital redlining. So just to level set, for every body, why don't we tell people first and remind people, what is regular redlining? And then we'll move into the digital space.

Chris Willis:

Sure. And the discussion of what I'll call traditional redlining, as we've understood it for more than a decade, is really exemplified by most recently, the Department of Justice's and CFPB's settlement with Trustmark Bank, which just came out a couple of weeks ago. And redlining historically has existed before even we had a CFPB, in the idea that if you had a mortgage lender, and in the past, they have always been banks, not non-banks. Although now, we have one non-bank case, the TownStone case brought by the CFPB. But traditionally, it's been banks. And the redlining theory is the idea, and the word red comes from the idea that you take a map of a city, for example, and you draw a red line around the neighborhoods where you don't want to make loans. And you would avoid making loans in those areas.

Chris Willis:

In fact, I think one of the very early cases involved someone taking a red marker to a map and drawing a red line. I think that's where it got the name. And so over the decades that followed in the hands of the Department of Justice and HUD, and the federal banking regulators, it developed as a claim either under the Fair Housing Act or the Equal Credit Opportunity Act or both, because remember, we're talking about mortgage lending, that the lender was purposely trying to avoid making loans in areas of a metro area that had a majority minority population. And so the Department of Justice, the federal banking regulators, and then later the CFPB, would go about proving these cases by essentially showing both the absence of lending activity in those majority minority neighborhoods and the absence of any attempt to do lending in those areas.

Chris Willis:

So advertising wasn't being directed there. Loan offices or branches were not located there. And none of the mortgage loan officers who were originating mortgage loans in the community were from there, or looked like the people who lived there, or things like that. And it's a type of intentional discrimination. It's always been understood to be that. And so it's been an area of pretty significant interest by the federal regulators during the Obama administration. You saw a lot of redlining cases being

brought by both Department of Justice and the CFPB. And then really, the CFPB kept on it during the years that Trump was president. The TownStone case, which was a very innovative redlining case, the first one ever brought against a non-bank, was brought during the Trump administration. And we saw the CFPB doing targeted redlining exams, the OCC was doing it.

Chris Willis:

So we saw all the federal regulators continue to be interested in it, but in a less public way, for the most part. And so now that we've had an administration change, the idea that the Department of Justice would now announce like, "We're doing a big initiative to combat redlining," is to me very expected, because it's always been an area of focus and it never went away even during the asked administration. And now, I think there's a desire and an ability to pursue it more vigorously. So it was really not a surprise to see the DOJ, the OCC, the CFPB jointly settled this case and agree to cooperate with one another in bringing more redlining cases. So that's the traditional understanding of redlining and the expected emphasis on it that we're going to see over the next several years from the federal regulators.

Alan Kaplinsky:

Yeah. And I guess, I mean that it's not a new concept at all. It's been with us as you described for a very long period of time. But now, a relatively new concept has been thrown into the mix. And that is Director Rohit Chopra of the CFPB is part of an announcement about a joint initiative with DOJ and the OCC. He made some comments about so-called "digital redlining" using algorithms. Tell us what Director Chopra was said. And then maybe you could translate what he said for those of us who are not as steeped in the technocratic lingo.

Chris Willis:

Sure. So as you noted, Director Chopra had a statement associated with the joint initiative with the DOJ and the OCC and the Trustmark settlement, which came out a couple of weeks ago. And in it, he reaffirmed the CFPB's commitment to combat what we just described as traditional redlining. But then he went on to say that equally, he's very concerned with the industry phenomenon in the consumer financial services industry of using algorithm to make credit underwriting decisions. He just used the word "algorithms," and we'll talk about what that means in a second. But he then went on to say that he is worried that the use of such algorithms can either perpetuate or worsen discrimination in lending decisions. And he called it a form of digital redlining.

Chris Willis:

And so what you have here, and this has been happening with the consumer advocates for the last few years, there's a desire to rebrand the term redlining to mean more than it has meant in the past. What you and I just talked about, the mortgage lending in a particular metro area, there's a desire to say, "Well, any practice by a lender that originates loans, say online, say a personal loan, or a student loan, or credit cards, or whatever, can be guilty of digital redlining if they do something that is discriminatory in terms of the manner in which the product is offered."

Chris Willis:

And so sometimes you'll hear digital redlining, I Don think it has an accepted, clear definition like traditional redlining definitely does. But I think sometimes you'll hear digital redlining used to refer to advertising practices like, "Oh, I advertised to this group and not that group. And so I'm redlining one group out of getting my product." In this instance, though, what the director of the bureau seems to be talking about is he's referring to algorithms. I think what he's talking about is newer technology algorithms that are more machine learning-based and may include alternative data elements. Because the lending industry's been using a kind of algorithm for 50 years, probably, to underwrite. Classic logistic regression automated underwriting systems have existed long enough that they're discussed in Regulation B, when which came out in the '70s.

Chris Willis:

And so I think what he's doing is he's responding to consumer advocates' suspicion that the use of machine learning algorithms, and particularly those that ingest and use alternative data beyond credit bureau attributes, may have some undesirable, discriminatory impact. And so therefore, he's terming that digital redlining. I think actually, that fear on the consumer advocates' behalf and the concern that the director is voicing is probably overblown because I think the new algorithms are better and less discriminatory than the old logistic progression ones. But we can talk about that in a minute, but I'm pretty sure that's what he's talking about, is he's picking up and echoing this consumer advocate suspicion of machine learning models and alternative data.

Alan Kaplinsky:

He didn't get into any detail, did he? Or he didn't give any ex any examples of what he was thinking about?

Chris Willis:

No, he just made some general statements. He just made the general statement that he's worried about algorithms, that they could be discriminatory, and that we should never assume that an algorithm is fair, essentially. And non-discriminatory in its application. It was statements of that general nature.

Alan Kaplinsky:

Well, okay. Put on your Rohit Chopra cap right now and see if you can conjure up the kind of thing he may be thinking about, a particular algorithm that maybe he has seen or heard about from the consumer advocates that prompted him to take this swipe at algorithms.

Chris Willis:

Sure. There's a few that I can point to for you. So for example, there have been two senators, Senator Warren and Senator Brown, who have done a lot of investigations, and public statements, and letters to the CFPB about the use of certain variables in student lending that are related to what school you go to. And so how well do students from your school repay their loans? What kind of income do they have when they get out? And therefore, that is used as a proxy for underwriting you for a student loan. And they've felt that the use of those school-specific variables is discriminatory, for example, against people who attend historically black colleges. And so they've urged the CFPB to take action with respect to the use of attributes like that in models.

Alan Kaplinsky:

That, sometimes, can referred to as cohort underwriting?

Chris Willis:

One of the variables is called the cohort default rate. It's a statistic maintained by the us Department of Education who uses it to qualify schools to receive or not receive Title IV student aid, but it's not okay in the consumer advocates and the senator's minds to use that in credit underwriting for student loans, which is a bit of an irony, I think. There was another very publicized case from, I don't know, 3, 4, 5 years ago where one of the large tech companies tried out a machine learning algorithm to try to screen applicants for particular jobs. And there, they allowed the machine learning algorithm to ingest all kinds of information, really in an unrestricted way about candidates for jobs, and then who was successful in the jobs. And so the algorithm had access to people's names, and birth dates, and other demographic information.

Chris Willis:

And so the algorithm basically decided that men were better for these jobs than women. And it started rejecting all the women. It was something like that. And the tech companies said they didn't actually use it and didn't put into operation, but it's used as a case study by opponents of machine learning of why you can't let the algorithm think for itself and train itself on

an unlimited source of data, which is definitely true, by the way. And creditors never do that, in my experience. And then I think there's this general idea that any machine learning algorithm has to be trained. It has to study a set of data and look for correlations between data points that a creditor can know at the time of an application and then later performance in terms of, "Do you pay your credit obligation or not?"

Chris Willis:

And so the belief among many consumer advocates is that the data sets themselves are biased because they may not represent members of protected classes as well, or they may even reflect sort of the historical impacts of discrimination and therefore be biased against members of protected classes. And so if you then do a completely fair and neutral training exercise of a machine learning algorithm on what they would refer to as a flawed data set, you get a discriminatory model, is sort of their view of it. Which again, some of that can be true. We can have data sets that are too limited. But if your idea is, "Hey, there's historical impacts of discrimination on members of protected classes," I don't know how you expect any algorithm to solve that problem if the point of the algorithm is to predict credit repayment performance. So there's a little bit of a question here about whether is this a violation of law or is it a social engineering exercise?

Alan Kaplinsky:

Yeah. So you said a few minutes ago, Chris, that you believe that the use of algorithms actually tends to eliminate discrimination. It's actually a good thing. Can you explain what you meant by that and maybe provide some examples of that?

Chris Willis:

Sure. So like in the old days from the 1970s until recently, the predominant form of automated underwriting that was used in the financial services industry was logistic regression models. And that's where you'd have modelers go, and with human minds, study the correlation between various input variables, like your credit score, or how many delinquencies you have, or what kind of trade lines you have, or what your credit utilization is or whatever. And then the outcome of how likely are you to repay a credit obligation. And so the logistic progression models are not that sophisticated. So you study the population as a whole, and then each attribute, like let's say, number of delinquencies in the past six months, then has a certain correlation with repayment or not repayment, that then is assigned a particular weight and a particular direction for everybody in the population.

Chris Willis:

So it treats everybody the same. And so it's a pretty blunt instrument, actually, for assessing individual consumers' likelihood of repaying a credit obligation. So those then tend to react to the data in a similar way to a machine learning model, except in a much less fine grained way, because they treat everybody the same. So now flip over and you do a machine learning model to underwrite for the same product. What it does is it doesn't just look at the relationship between a single variable and the outcome. So say, delinquencies versus your repayment of a new credit obligation. But it looks at the combination of, "Well, if you have this many delinquencies in the last six months, but you also of this credit utilization, and you also have this number of inquiries," and it groups variables together and looks for associations, not with them individually, but with combinations of them together.

Chris Willis:

And so it ends up, one of the most common machine learning model building techniques is called a boosted tree or a gradient boosted tree, and it creates this forest, this forest of trees with all kinds of thousands of different scenarios of combinations of variables together. And it gives the variables different weights or even different directions depending on their association with other variables. And so it tends to treat consumers in a much more personal way, and not everybody gets the same one-size-fit-all. So as a result of that, even without trying, a lot of machine learning models tend to be more inclusive and give greater approvals to members of protected classes than your traditional logistic aggression models. And then in addition, if you are able to use alternative data points that will allow you to score someone who doesn't have a traditional credit bureau report

that's sufficiently thick to give rise to a credit score, then you can make credit available to a lot of members of that group who would get declined if you're just using traditional credit bureau data.

Chris Willis:

And in fact, the CFPB has issued three reports on what it refers to "credit invisibles," that is people who are thin files or no hits with the three big credit bureaus, and saying that those are predominantly members of protected classes. And that the use of alternative data, things that are not on a traditional credit bureau report, is really the way to make credit available to them. And so machine learning models do a better job of ingesting and dealing with the volume of alternative data than a traditional logistic regression model does. And so they can create a lot of credit inclusiveness to those thin file and no hit credit invisible population, which again, is beneficial members of protected classes who might be frozen out of the credit market if we just did things the old way.

Alan Kaplinsky:

Yeah. You wonder why... It's obvious Rohit is a very bright guy. What do you think is going on here? Is it just careless use of a throw away line or do you think it was very carefully thought through, the way you have done today?

Chris Willis:

I don't think it's careless. I don't impute carelessness to him, because as you noted, he's very bright and very thoughtful. I believe he is responding to the suspicion on the consumer advocacy side, that there are models that are built using these new technologies that can be discriminatory. And that's definitely true. If you do it the wrong way, you for sure can get a discriminatory model, but you also can do it the right way. And his comments were very weighted towards, "Hey, I think it's being done the wrong way." And the thing that was really, to me, very strange, was the idea that it's bad to use quote algorithms. And there, you throw back to a debate between judgemental underwriting, like having a human being, make an underwriting decision, and having a logistic regression model, like the industry has for the past however many 50 years, and we've consistently seen fair lending concerns motivate regulators to push lenders into greater use of automated underwriting models where there's no human discretion, there's no judgment.

Chris Willis:

And so the view has been that the extra size of judgment and discretion has always been a vehicle for disparate impact. And how many discretionary underwriting and discretionary pricing cases were there? In mortgage first, and then you saw them in auto finance too. And so are we seriously saying that automated underwriting should be done away with in favor of more judgmental underwriting? Because that doesn't seem sensible to me, because we've spent 20 or 30 years with the federal regulators fighting against that and wanting more automated underwriting. And so that was puzzling to me.

Chris Willis:

And then there was also a reference in his comments in his recent testimony in Congress to wanting to move the industry back towards relationship banking, which to me, I don't know what he means by relationship banking. But to me, it hearkens back to the idea of, you have a local bank branch and they know you because you come in there all the time, and you make your deposits and your withdrawals there, and they know about your employment, and credit, and your character. And so they'll make lending decisions based on their relationship with you, which first of all, that model doesn't really seem all that applicable to modern times where people don't really go into bank branches that much anymore.

Chris Willis:

But even if it were, again, that sort of, "I'm going to treat my friends preferentially," or "I'm going to people I already know preferentially," seems likely to me to be another avenue for suspected discrimination, along the same judgmental discretionary lines that we just talked about. So surely, I don't think we're going to see the CFPB say, "Oh, everybody throw away your automated underwriting systems and do judgmental again." That's, I don't think going to happen, but it's part of reason that

it's difficult to understand the recent comments, because it does seem to romanticize the old way of doing things that, I think in a lot of the federal regulators' minds, was the bad old way.

Alan Kaplinsky:

Yeah. Reminded me of, well, It's a Good Life, I think it's called, the movie. I got the title wrong, I'm sure. With Spencer Tracy or who managed this little community bank. I mean, that's sort of what I thought it sounded like.

Chris Willis:

Yeah, exactly.

Alan Kaplinsky:

Yeah. So this is very frustrating because what do you do, given a comment like the one that Director Chopra made, what do you do in advising clients about using digital algorithms in their underwriting? Do you ignore that comment and you say, "Listen, as long as you don't use something that results in no loans being given to men, or women, or members of a particular race or particular age, you're going to be okay." or do you tell them to tread much more carefully than that?

Chris Willis:

Yeah, so I'm really honestly really worried that the director's comments are going to be seen as being very hostile to the use of machine learning and alternative data by the industry. And I think that that's bad for the industry and it's bad for consumers, because the machine learning models are generally much better. They're more accurate, they're better at evaluating credit risk, so they allow better underwriting and pricing by financial institutions. And then as I said, they allow for inclusion of people who might be frozen out if you use the old kind of model using just traditional credit bureau data.

Chris Willis:

And I'm really worried that actually, the industry will take these kinds of public comments as hostility to this kind of innovation and say, "Well, okay, the CFPB doesn't want me to do that. I'll just stick with the old way, essentially." Personally, I don't think that is warranted from an industry standpoint, because the new models can be built and governed in a way that is far superior from a fair lending standpoint than the old way. And so it seems counterintuitive for the industry to say, "Well, here's this new technological innovation that will allow me to be more profitable, and allow me to approve more loans to members of protected classes, but I'm not going to do it because of these public statements by the CFPB director."

Chris Willis:

That doesn't make sense to me. That should be the reaction of the industry. But at the same time, you can't just ignore it. The CFPB's director is sending us a super clear signal that this is going to be an area of significant scrutiny. So what do you do? From my standpoint, I don't think you turn your back on the technology and the benefits of it. And I think market forces would really preclude us from doing that anyway. But I think what you do is you really, really focus on making your model development and testing exercise, one, that can provide abundant evidence of both the business justification for the model, that is how accurate it is in predicting repayment performance. And the fact that adopting the model allowed you to approve more members of protected classes, which you can easily prove. And then in addition to that, lenders can investigate some of the new techniques for trying to find the least discriminatory versions of their models.

Chris Willis:

There's a couple of technologies available on the market now for, in an automated way, making changes to a model and tuning it in a way to minimize disparate impact from the model. Even if the model has a very strong business justification. And so pursuing something like that, one of these de-biasing techniques, I think, is also a very good idea. So to me, financial institutions should get ready to defend their models and be able to prove with data that the model was built in a sound way and that it has a good outcome from a fair length standpoint. But I don't think they should just say, "Forget it. I'm not doing

machine learning. I'm not doing alternative data. This is too hot for me in this kitchen," because again, both the consumer benefits and the market benefits of it are just too great, I think, to throw your hands up and say, "I'm not doing it."

Alan Kaplinsky:

Yeah, yeah. It's very strange this, I mean, what we're talking today, about digital redlining and one comment made by Director Chopra. But he's done a number of things, both as a commissioner of the FTC and now as director of the CFPB, that could be viewed as sort of anti-technology, being skepticism, I guess, of technology. And almost saying that not giving the benefit of the doubt to FinTech, and rather, saying to those that are employing FinTech, that "You've got to prove to me that this is better than the old fashion method."

Chris Willis:

In fact, he explicitly said that in his comments associated with the Trustmark settlement. He specifically said, "We should never assume algorithms are fair." It may not be word for word, but he said something very similar to that. So your statement that the burden of proof is on the entity using the technology to prove that it is fair rather than the burden being on the regulator to prove that it's a violation of law, you're right on. And the industry needs to get ready for that and get ready to shoulder that burden of proof, I think.

Alan Kaplinsky:

Yeah. I mean, I think the thing that's gotten all the media attention lately is the broad side that he made to a number the major technology companies, Apple, Facebook, Google. I think they've all been the subject of requests for information dealing with their privacy practices and how they maintain data, how they use it, et cetera, et cetera. And nobody really knows exactly where he is going, but it almost seemed like he just wanted to say, "I'm going to give you guys a tough time. You're not going to have an easy ride with me."

Chris Willis:

I think it's fair to say that the hostility seems to be evident.

Alan Kaplinsky:

Yeah. Yeah, for sure. So before we wrap it up today, Chris, we are drawing to the end of our program. Do you have anything else regarding this topic of digital redlining that you'd like to share with our listeners?

Chris Willis:

All I would say is if you're in the industry, and that's a lot of our audience, pay careful attention to what you can do to both build a model the right way and be in a position to prove that you did so, and that the outcomes of the model are fair, and fairer than the alternative of what you were using before or what you might otherwise use. And I think doing that and being prepared to prove it is a necessity in today's regulatory environment. And then for the FinTech firms that are out there who are developing these technologies, they should remain engaged with the CFPB.

Chris Willis:

A lot of them have been very engaged with the bureau, and it's innovation office, and it's fair lending office. To illustrate how these things can work and how they can be beneficial to consumers. And I think that continued engagement with the bureau's staff, for example, in the fair lending office, I think is very important. Because I think the folks in the fair lending office have a very sophisticated view of this, and a good understanding that it's a technology that can be very helpful to consumers, and helpful in helping reverse sort of the legacy of discrimination in society, and that over deterring it is probably not the way to go.

Chris Willis:

So I think continued engagement with the bureau to continue to illustrate that, I think, is necessary.

Alan Kaplinsky:

Yeah. And as you say, don't forget about their office of technology, because I would think as a general matter, that office is trying to support the use of technology in consumer finance, and that they may very well be a good ally in any arguments that the industry may have with the director himself.

Chris Willis:

Yeah. Could be. But certainly, keeping them in the loop to promote a balanced view of these issues, I think, is necessary and advisable.

Alan Kaplinsky:

Well, Chris, thank you very much for sharing your thoughts today. Also, really want to thank all of our listeners who have downloaded the show today and remind all of you that our show is available weekly, except for two weeks where we go on vacation. Thanksgiving week and Christmas week. We release a new show every Thursday other than during those two weeks. Thank you again.