

February 7, 2020

HUD, the federal agency that provides housing assistance to over 4 million American families, is proposing to give landlords a legal safe harbor to discriminate against tenants. They recently put forth a new rule that, as long as landlords rely on algorithms for tenant screening, they would bear no responsibility for any discriminatory results. While HUD should indeed bring its practices in line with landlords' current use of big data and algorithms, the proposed rule is worse than doing nothing.

The issue is compliance with the Fair Housing Act, which prohibits discrimination in housing on the basis of race, religion, or any other legally protected characteristic. This law applies whether landlords review applications by hand or, as is common today, use AI tools and algorithms to assess the quality and riskiness of tenants. Tenants who believe they have been unfairly denied by a landlord can make a discrimination claim. One type of discrimination claim is a "disparate impact" claim, which applies when the landlord's selection process has a disproportionate impact on a particular race, gender, or other protected class. HUD's proposed new rule give landlords who use algorithms three ways to defeat such claims: by showing that protected classes or "substitutes" for such classes are not actually used in the algorithm, by showing the algorithm was built and maintained by a "recognized third party," or by having a neutral auditor certify that the algorithm is, among other things, "demonstrably and statistically sound" (whatever that means; HUD does not define it).

I am an unlikely critic. I founded my company ORCAA specifically to conduct algorithmic audits. The third proposed rule, then, would seem to be a boon to me and my company. But the utter vagueness of HUD's proposal shows they are interested in creating a loophole for bad behavior, not meaningful standards for good behavior. The "audit" defense offered to landlords does not define who is qualified to conduct an audit, what the process entails, or how often it should happen, nor does it address that a model that is statistically sound could still be discriminatory or how the statistical soundness of a model has any relation to whether it is discriminatory. The predictable result would be a race to the bottom: "neutral third parties" happy to rubber-stamp any algorithm as "sound," for a price. I'd rather my chosen industry have good standards.

The second proposed defense gives landlords an even easier way to deflect discrimination claims: instead of developing your own tenant screening algorithm and having it audited, just use one built by a "recognized third party" (again undefined) and you will never have to answer for it. This policy would create terrible incentives for landlords. Unaccountable for the algorithm's decisions, they would have no reason to find out whether a particular screening algorithm is fair in the first place, and no reason to pay attention to complaints once they decide to use it.

The first proposed defense, defending the variables used in the algorithm, is also deeply misguided. It is tempting to think we could make an algorithm "race blind" by excluding race variables from the data

altogether. We know it's not that easy. Researchers from Harvard and Microsoft found that removing race data from a mortgage application dataset did not fix racial disparities in algorithmic pricing; in fact it widened the gap in some cases. The machine learning techniques behind these algorithms can effectively reconstruct missing variables by combining other correlated variables in complex ways. It is not enough to require that certain variables be left out—the results could still be unfair. In fact, given that historical data will naturally reflect the practice of redlining we know we've been doing, we should expect the result to be unfair along racial lines.

In sum, all of HUD's proposed defenses give landlords a way to deflect disparate impact claims without addressing the issues. Beyond these defenses, the new rule also raises the burden of proof required for tenants to bring a disparate impact claim. Why does HUD want rules that make it easier to dodge discrimination claims and harder to bring them?

There's a better approach. If HUD wants to look into landlords' use of algorithms for screening tenants -- and it should -- then it should stick to the purpose behind the established legal concept of disparate impact. According to the Supreme Court, disparate impact discrimination occurs when a seemingly neutral policy produces an unjustifiable discriminatory effect. Critically, it focuses on the effects, rather than the intentions, of housing policies. The question HUD should be asking is whether there are disparities in who actually gets housing assistance. If HUD wants to make progress against unfair treatment, they should establish specific thresholds for allowable gaps between rejection rates for different groups to be compliant with the Fair Housing Act. If they don't want to do that, they should step away from the question altogether.

