

7-2-2020

Reconciling Risk and Equality

Christopher Slobogin

Follow this and additional works at: <https://scholarship.law.vanderbilt.edu/faculty-publications>



Part of the [Civil Rights and Discrimination Commons](#), [Constitutional Law Commons](#), and the [Law and Society Commons](#)

Recommended Citation

Christopher Slobogin, *Reconciling Risk and Equality* Jotwell. (2020)

Available at: <https://scholarship.law.vanderbilt.edu/faculty-publications/1172>

This Article is brought to you for free and open access by the Faculty Scholarship at Scholarship@Vanderbilt Law. It has been accepted for inclusion in Vanderbilt Law School Faculty Publications by an authorized administrator of Scholarship@Vanderbilt Law. For more information, please contact mark.j.williams@vanderbilt.edu.

Reconciling Risk and Equality

Author : Christopher Slobogin

Date : July 2, 2020

Deborah Hellman, *Sex, Causation and Algorithms: Equal Protection in the Age of Machine Learning*, 98 **Wash. L. Rev.** ___ (forthcoming, 2020), available at [SSRN](#).

States have increasingly resorted to statistically-derived risk algorithms to determine when diversion from prison should occur, whether sentences should be enhanced, and the level of security and treatment a prisoner requires. The federal government has jumped on the bandwagon in a big way with the First Step Act,¹ which mandated that a risk assessment instrument be developed to determine which prisoners can be released early on parole. Policymakers are turning to these algorithms because they are thought to be more accurate and less biased than judges and correctional officials, making them useful tools for reducing prison populations through identification of low risk individuals.

These assumptions about the benefits of risk assessment tools are all contested. But critics also argue that, even if these instruments improve overall accuracy, they are constitutionally suspect. While no instrument explicitly uses race as a “risk factor” (which in any event is probably barred by the Supreme Court’s decision in *Buck v. Davis*²), several do incorporate sex (with maleness increasing the risk score) and many rely on factors that are highly correlated with race or socio-economic status, which is said to violate equal protection principles.³

In *Sex, Causation and Algorithms*, [Deborah Hellman](#), a philosopher and constitutional law scholar, provides some provocative food for thought on this issue. The article focuses on the Supreme Court’s Fourteenth Amendment caselaw on sex as a classification. But the approach to equal protection that Hellman develops could also provide a response to many of the other discrimination and disparate impact challenges aimed at risk assessment instruments.

Hellman proposes what she calls an “anti-compounding injustice” theory of equal protection, which presumptively prohibits use of sex as a classification when, but only when, the classification would “compound” sex-based injustice. For instance, while she agrees with the Supreme Court’s decision in *Frontiero v. Richardson*,⁴ which held that the military may not use sex as a proxy for whether a member of the service is a financially dependent spouse, she would adopt a different rationale. The Court looked at the “fit” between sex and spousal dependency (which, at the time, was fairly good, and thus did not obviously support the Court’s conclusion). Hellman would instead look at whether the fact that women tend to be the dependent spouse was the result of sex-based injustice (which it was, given society’s longtime privileging of male domination). Hellman’s analysis of *United States v. Virginia*⁵ is similar. While the Court in that case held that denying women admission to a military college would reaffirm demeaning stereotypes even if only one woman had the “will and capacity” to enter, Hellman points out that this type of reasoning is, in effect, strict scrutiny, not the more relaxed intermediate scrutiny supposedly applicable in sex cases. Hellman argues that the better rationale for the case is that preventing women from entering the academy would be compounding the sex-based injustice that has made women less likely to be willing and qualified to enter such schools.

Hellman contends that her anti-compounding injustice theory is consistent with most of the Court’s cases, at the same time it is less confusing than the Court’s current focus on whether a sex classification closely fits the state’s goals, exacerbates stereotypes, or reflects “real differences” between the sexes. She also argues that her approach is more morally compelling. She is persuasive on both points. But what does this have to do with risk algorithms?

Hellman starts the paper with a reference to *Wisconsin v. Loomis*,⁶ where the defendant was sentenced to six years after the trial court considered testimony based on the COMPAS, a risk algorithm. Loomis argued that the use of the COMPAS violated due process, because it considers sex as a risk factor. The Wisconsin Supreme Court demurred,

noting that “any risk assessment tool which fails to differentiate between men and women will misclassify both genders;” in other words, as an empirical matter, if sex is not taken into account, a woman whose risk factors are otherwise identical to a man’s will be rated as higher risk than she actually is. While Loomis lost his due process argument on accuracy grounds, he might well have won had he framed his challenge in equal protection terms, because the algorithm’s explicit consideration of sex violates the Supreme Court’s current anti-classification approach to the Fourteenth Amendment.

Hellman suggests that result would be wrong. Taking sex out of the algorithm would harm women, which would be compounding injustice, because “the bulk of gender-based injustice has harmed women.” Although Hellman doesn’t explicitly say so, she likewise seems to disagree with the outcome of *Craig v. Boren*,⁷ where the Court struck down a statute that set the drinking age at 21 for males and 18 for females, despite evidence that men were 10 times more likely to drink and drive than women. Her approach thus seems akin to anti-subordination theory.

That observation raises the question of how Hellman would treat disparate impact challenges against algorithms. Many algorithms use risk factors that correlate with race, including employment status, location, and criminal history, the latter the predominant risk factor in every risk algorithm. Current equal protection doctrine would consider such correlations irrelevant unless intent to discriminate can be shown. While Hellman does not directly address this issue, she suggests that anti-compounding theory would approach these cases differently and could even be “revisionary.” She poses a hypothetical in which the state enhances the sentence of an offender who was abused as a child because an algorithm indicates that child abuse is a risk factor, and argues that, regardless of whether discriminatory intent is present, her anti-compounding injustice theory would call such an algorithm into question. One might make the same argument against including a risk factor such as unemployment or criminal history on a risk tool if it correlates with race, given the likelihood that unemployment and criminal offending are higher among people of color because of race-based injustice.

However, Hellman also says this: “Compounding injustice is not a decisive reason to avoid an action in all contexts, nor is the duty to avoid such compounding injustice a duty that trumps everything else,” and adds, in connection with the victimization hypothetical, “[t]he interests of other people – those whom [the individual] may harm if he is released – count as well.” In other words, as is true under traditional equal protection theory, a strong state interest can override compounded injustice. As applied to risk algorithms, this caveat might mean that use of a risk factor correlated with race is permissible if it is a robust predictor. That may not be true of unemployment status, but it is certainly true of criminal history.

Risk algorithms surface a real tension between traditional equal protection law and the goal of ensuring that predictions are as accurate as possible (a tension that exists, by the way, whether prediction is based on algorithms *or* on subjective judgments, which rely on the same factors as algorithms, only more opaquely so). Hellman’s anti-compounding theory may help courts and criminal justice scholars figure out how that tension should be resolved.

1. 18 U.S.C. § 3261 (2018).
2. Sonja B. Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, 66 **Stan. L. Rev.** 803 (2014).
3. 137 S.Ct.. 759 (2017).
4. 411 U.S. 677 (1973).
5. 518 U.S. 515 (1996).
6. 881 N.W.2d 749 (Wisc. 2016).
7. 429 U.S. 190 (1976).

Cite as: Christopher Slobogin, *Reconciling Risk and Equality*, JOTWELL (July 2, 2020) (reviewing Deborah Hellman, *Sex, Causation and Algorithms: Equal Protection in the Age of Machine Learning*, 98 **Wash. L. Rev.** __ (forthcoming, 2020), available at SSRN), <https://crim.jotwell.com/reconciling-risk-and-equality/>.