PRIVATIZING SENTENCING: A DELEGATION FRAMEWORK FOR RECIDIVISM RISK ASSESSMENT

Andrea Nishi*

As the use of predictive technology expands, an increasing number of states have passed legislation encouraging or requiring judges to incorporate recidivism risk assessment algorithms into their bail, parole, and sentencing determinations. And while these tools promise to reduce prison overcrowding, decrease recidivism, and combat racial bias, critics have identified a number of potential constitutional issues that stem from the use of these algorithms. Because state governments often contract with private companies to develop and license these tools, defendants and judges have limited information about the development and operation of the predictive software at the heart of these constitutional claims, raising a number of questions about the strength of constitutional protections when private actors are involved in the sentencing process.

This Note addresses the inadvertent and largely unexamined role that private actors—risk assessment developers—have come to play in individual sentencing determinations. The private nature of many risk assessment algorithms leaves sentencing judges unable to understand and adequately apply algorithmic results, leading to a greater reliance on the undisclosed policy decisions of private developers. Arguing that this has given private actors an outsized role in sentencing, this Note proposes a solution for increasing accountability and legitimacy in recidivism risk assessment based on case law addressing delegations to private actors. Because existing statutory frameworks allow private actors to wield government power with limited public oversight and control, legislators must strengthen recidivism risk assessment statutes, increasing the ability of judges to understand and apply algorithmic risk scores.

INTRODUCTION

In a 2017 interview, Chief Justice John Roberts was asked if he could foresee a day when artificial intelligence would play a role in judicial decisionmaking. Immediately noting that "[i]t's a day that's here," the Chief Justice described several types of technology that are widely used in the judicial system, concluding that the courts "have not yet really

absorbed how it's going to change the way we do business."[1] Roberts was right—courts today make use of a variety of new technologies, including artificial intelligence,[2] but the impact of these innovations on the judicial system remains largely unaddressed.

Artificial intelligence already plays a significant role in judicial decisionmaking through the widespread use of recidivism risk assessment algorithms in state criminal justice systems. Today, at least twenty states use risk assessment algorithms to predict recidivism in their bail, parole, and sentencing proceedings,[3] encouraging or requiring judges to consider them in making their determinations.[4] And while these tools promise to decrease recidivism, reduce prison overcrowding, and combat racial bias,[5] critics have identified a number of potential constitutional issues that stem from the use of these algorithms, including due process and equal protection claims.[6]

Further complicating these constitutional questions is the fact that state governments often contract with private companies to develop and license these tools, in part due to the level of technological and statistical expertise needed to create them.[7] Because of this outsourcing, defendants and judges alike have limited information about how these risk assessment systems operate. For example, in a 2016 case from Wisconsin, a private developer denied an offender's request for information about the algorithm that was used to determine his sentence on the ground that it was a trade secret.[8] As states increasingly contract with private companies to incorporate algorithmic risk assessment into their criminal

---

1. Rensselaer Polytechnic Inst., A Conversation with Chief Justice John G. Roberts, Jr., YouTube (Apr. 12, 2017), https://www.youtube.com/watch?v=TuZEKlRgDEg (on file with the *Columbia Law Review*).

2. Artificial intelligence can be defined broadly as "a branch of computer science dealing with the simulation of intelligent behavior in computers," Artificial Intelligence, Merriam-Webster, https://www.merriam-webster.com/dictionary/artificial%20intelligence [https://perma.cc/E5X8-NLDL] (last visited July 24, 2019), and can include a variety of different "intelligent" functions performed by machines, such as computational creativity, natural language processing, and machine learning. This Note only addresses machine learning, the ability of computers to improve their performance on an assigned task through data processing and repetition. See Ethem Alpaydin, Introduction to Machine Learning 3–8 (3d ed. 2014).

3. See Sonja B. Starr, Evidence-Based Sentencing and the Scientific Rationalization of Discrimination, 66 Stan. L. Rev. 803, 809 (2014).

4. See infra section I.A.

5. See infra section I.A.

6. See infra section I.C.

7. See Cary Coglianese & David Lehr, Transparency and Algorithmic Governance, 71 Admin. L. Rev. 1, 30 (2019).

8. See State v. Loomis, 881 N.W.2d 749, 761 (Wis. 2016). The Wisconsin Supreme Court found no due process violation in the use of this proprietary software, in part because the sentencing judge had access to no more information about the risk assessment tool than the defendant did. Id.

justice systems,[9] a number of questions arise about the strength of constitutional protections when private actors play a role in the sentencing process.

While scholarship on risk assessment algorithms has focused primarily on the constitutionality of risk assessment and challenges to the invocation of trade secret protections,[10] this Note addresses the inadvertent and largely unexamined role that private actors—risk assessment developers—have come to play in individual sentencing determinations. Arguing that a lack of oversight and control by state actors has created an accountability gap in the use of these tools, this Note proposes a solution for increasing accountability and legitimacy in recidivism risk assessment based on case law addressing delegations to private actors. Part I provides a brief history of the development and use of risk assessment instruments in sentencing, followed by a more comprehensive explanation of how modern risk assessment tools differ from earlier models. Part II explains how the process for developing modern risk assessment algorithms has obscured the way these tools operate, making judges more likely to rely on these seemingly objective yet opaque assessments in sentencing decisions. This has inadvertently allowed private developers to play a significant role in sentencing individual defendants, while remaining unrestricted by traditional notions of constitutional accountability that bind state actors. Part III suggests that the private delegation doctrine—a largely dormant, New Deal–era doctrine developed to increase government oversight and control of private actors exercising government power—can provide a framework for understanding and filling this accountability gap.[11] The Note concludes in Part IV, which uses private delegation principles identified in Part III to craft legislative remedies to restore constitutional

---

9. See supra note 7, infra notes 22–26 and accompanying text.

10. See infra sections I.C, II.A.

11. Rather than focusing on the "black box" nature of risk assessment algorithms as the source of these accountability issues, this Note uses existing discussions of algorithmic opacity as a starting point for examining how these problems mask the role of private algorithm developers in sentencing decisions. This Note examines the use of risk assessment algorithms as a delegation of power to the individual human actors that develop those algorithms, whereas existing literature applying delegation concepts to the use of algorithms in government typically focuses on the delegation of government functions to algorithms themselves. See Mariano-Florentino Cuéllar, Cyberdelegation and the Administrative State 2 (Stanford Pub. Law Working Paper No. 2754385, 2016), https://ssrn.com/abstract=2754385 (on file with the *Columbia Law Review*) ("I consider here some of the trade-offs associated with the delegation of agency decisions to computer programs . . . ."); see also Cary Coglianese & David Lehr, Regulating by Robot: Administrative Decision Making in the Machine-Learning Era, 105 Geo. L.J. 1147, 1178 (2017) [hereinafter Coglianese & Lehr, Regulating by Robot] ("As machine learning becomes more advanced and government agencies use it more extensively, decision-making authority could effectively become delegated still further—to computerized algorithms.").

accountability to the use of privately developed risk assessment algorithms in sentencing.

I. RISK ASSESSMENT IN SENTENCING

This Part provides an overview of the history and current applications of recidivism risk assessment in sentencing. Section I.A describes the development of risk assessment tools in the twentieth century, as well as the policy reasons for using them, before looking at examples of state statutes that authorize or mandate risk assessment in sentencing. Section I.B discusses the machine learning process, illustrating how today's risk assessment algorithms differ from earlier risk assessment methodologies.[12] Finally, section I.C summarizes some of the constitutional concerns surrounding the use of modern risk assessment algorithms in sentencing, setting the stage for Part II's discussion of opacity, privatization, and constitutional accountability.

A. *The History and Use of Recidivism Risk Assessment Tools*

Now ubiquitous,[13] the use of risk assessment tools in the U.S. criminal justice system dates back to the 1920s, when sociologist Ernest Burgess developed a statistical method for estimating recidivism risk for parole determinations.[14] In the 1960s and 1970s, statisticians made further efforts to create models for identifying offenders with a high risk of committing violent crimes.[15] These early methods were not particularly

---

12. While recidivism risk assessment algorithms can be developed using a variety of different techniques, only one of which is machine learning, a basic description of how machine learning works is useful for understanding the inscrutability of algorithmic outputs, as well as the consequential policy decisions that private developers make when creating these algorithms, both of which are discussed in Part II. Due to the proprietary nature of many recidivism risk assessment tools, it can be difficult to identify the process used to develop a specific instrument. See infra note 74. Regardless of the methodology used to create a particular tool, the machine learning process provides a rich illustration of the complexity and subjectivity of developing a risk assessment tool using sophisticated modeling techniques, which is likely to become increasingly relevant as the use of machine learning continues to expand. See AI: Algorithms and Justice, Berkman Klein Ctr. for Internet & Soc'y at Harvard Univ., https://cyber.harvard.edu/projects/ai-algorithms-and-justice [https://perma.cc/VLJ5-CNY6] (last visited Aug. 16, 2019) ("Use cases for technologies that incorporate AI or machine learning will expand as governments and companies amass larger quantities of data and analytical tools become more powerful."); see also infra note 35 and accompanying text.

13. See infra notes 22–25.

14. See Alyssa M. Carlson, Note, The Need for Transparency in the Age of Predictive Sentencing Algorithms, 103 Iowa L. Rev. 303, 308 (2017); John Monahan, Anne L. Metz & Brandon L. Garrett, Judicial Appraisals of Risk Assessment in Sentencing 3 (Univ. of Va. Sch. of Law, Public Law & Legal Theory Research Paper No. 2018-27, 2018), https://ssrn.com/abstract=3168644 (on file with the *Columbia Law Review*).

15. Danielle Kehl, Priscilla Guo & Samuel Kessler, Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing 3–4 (2017), https://dash.harvard.edu

accurate or useful, with some identifying up to ninety-nine percent of study participants as "dangerous."[16] Perhaps because of the unreliability of these early models, clinical risk assessments, in which correctional staff and clinical psychologists would undertake unstructured, one-on-one interviews to assess an individual's likelihood of recidivating, were more popular during this time.[17] In sentencing specifically, risk assessment became common in the 1980s, when sentencing commissions began to use criminal history as an approximation of recidivism risk in sentencing guidelines.[18]

Although the instruments for assessing risk have changed since these tools were first developed, three policy arguments continue to drive the use of risk assessment tools. First, risk assessments may help to reduce prison populations and save taxpayer money by enabling judges to sentence low-risk defendants to shorter prison terms.[19] Second, they increase fairness in the criminal justice system by providing an assessment of a defendant's dangerousness, purportedly free from bias that may plague judicial decisionmaking.[20] Finally, they reduce recidivism

---

/bitstream/handle/1/33746041/2017-07_responsivecommunities_2.pdf [https://perma.cc /K9LE-VLUJ].

16. Id. at 4.

17. See, e.g., id. at 4–6; Jessica M. Eaglin, Constructing Recidivism Risk, 67 Emory L.J. 59, 67 n.34 (2017).

18. See Eaglin, supra note 17, at 67.

19. See Starr, supra note 3, at 816; see also Monahan et al., supra note 14, at 3 ("[O]ne way to begin dialing down 'mass incarceration' without simultaneously jeopardizing the historically low crime rate is to put risk assessment back into sentencing."); cf. Rick Jones, The Siren Song of Objectivity: Risk Assessment Tools and Racial Disparity, Champion (Apr. 2018), https://www.nacdl.org/Champion.aspx?id=52177 [https://perma.cc/Z42Z-52WF] ("Theoretically, RAIs help the criminal justice system fulfill its 14th Amendment mandate of due process by reducing overly burdensome bail requirements and lengthy periods of pretrial detention.").

20. See Matthew DeMichele, Peter Baumgartner, Kelle Barrick, Megan Comfort, Samuel Scaggs & Shilpi Misra, What Do Criminal Justice Professionals Think About Risk Assessment at Pretrial? 3 (Apr. 25, 2018) (unpublished manuscript), https://ssrn.com/ abstract=3168490 (on file with the *Columbia Law Review*); see also Filippo A. Raso, Hannah Hilligoss, Vivek Krishnamurthy, Christopher Bavitz & Levin Kim, Berkman Klein Ctr. for Internet & Soc'y at Harvard Univ., Artificial Intelligence & Human Rights: Opportunities & Risks 23 (2018), https://cyber.harvard.edu/sites/default/files/2018-09/2018-09_ AIHumanRightsSmall.pdf? [https://perma.cc/674G-LQ7W] (describing a study finding that New York City "could reduce its crime rate by 25% by incarcerating the same number of people, but changing the criteria for who gets bail . . . with concomitant positive effects on the right to equality and non-discrimination"). Rampant racial bias has nevertheless been documented in the results of risk assessment algorithms. See Julia Angwin, Jeff Larson, Surya Mattu & Lauren Kirchner, Machine Bias, ProPublica (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing [https://perma.cc/G338-U2ML]; see also Jones, supra note 19.

and increase public safety by enabling judges to better understand a defendant's rehabilitative needs.[21]

The use of predictive risk assessment tools in sentencing has increased dramatically in recent years.[22] The first actuarial risk assessment instrument for sentencing was implemented in Virginia in 1994,[23] and as of 2014 approximately twenty states used some form of risk assessment in sentencing.[24] In the current working draft of the Model Penal Code, the American Law Institute has even added a provision encouraging states to incorporate risk assessment tools into the sentencing process.[25] In states that have adopted predictive risk assessments, an offender's risk score—which often appears in the form of both quantitative and qualitative "high," "medium," or "low risk" values—is typically provided to the sentencing judge in the offender's presentence investigation report.[26]

Many state legislatures have recently started to require the consideration of recidivism risk assessments during the sentencing process.[27] For example, Kentucky's state sentencing policy mandates that judges consider "the results of a defendant's risk and needs assessment included in the presentence investigation" and "[t]he likely impact of a

---

21. See Erin Collins, Punishing Risk, 107 Geo. L.J. 57, 72–74 (2018); see also Cassie Deskus, Note, Fifth Amendment Limitations on Criminal Algorithmic Decision-Making, 21 N.Y.U. J. Legis. & Pub. Pol'y 237, 243 (2018) ("When used correctly, these algorithms identify groups of defendants whose incarceration is likely to be unnecessary and provide them with alternative sources of rehabilitation, such as community supervision."). The use of recidivism risk assessment tools has become increasingly popular over the last twenty years as criminal justice policy has shifted in focus from retributivism to rehabilitation and the individual risks and needs of particular offenders. See Brandon L. Garrett & John Monahan, Judging Risk, 108 Calif. L. Rev. (forthcoming 2019) (manuscript at 6), https://ssrn.com/abstract=3190403 (on file with the *Columbia Law Review*).

22. See Kehl et al., supra note 15, at 14; see also Recent Case, *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016), 130 Harv. L. Rev. 1530, 1530 (2017).

23. Kehl et al., supra note 15, at 10.

24. See Monahan et al., supra note 14, at 3.

25. See Model Penal Code § 6B.09(2) (Am. Law Inst., Tentative Draft No. 2, 2011) (providing that state commissions "shall develop actuarial instruments or processes, supported by current and ongoing recidivism research, that will estimate the relative risks that individual offenders pose to public safety").

26. See Kehl et al., supra note 15, at 8, 15; see also Carlson, supra note 14, at 315. For an example of the risk assessment report that a judge might receive, see Sharon Lansing, N.Y. State Div. of Criminal Justice Servs., New York State COMPAS-Probation Risk and Need Assessment Study: Examining the Recidivism Scale's Effectiveness and Predictive Accuracy app. at 29 (2012), https://www.criminaljustice.ny.gov/crimnet/ojsa/opca/compas_probation_report_2012.pdf [https://perma.cc/H6EC-J5JP]. Additional information included in a presentence investigation report can include offender interviews, statements from victims, and information on the offender's mental health and criminal history. See Wash. State Dep't of Corr., Pre-Sentence Investigations and Risk Assessment Reports Ordered by the Court 2–3 (2014), https://www.doc.wa.gov/information/policies/files/320010.pdf [https://perma.cc/MXS7-DV2Z].

27. Eaglin, supra note 17, at 114.

potential sentence on the reduction of the defendant's potential future criminal behavior."[28] Likewise, Tennessee law mandates that "[i]n preparing presentence reports . . . the department of correction shall include information identifying the defendant's risks and needs as determined through the use of a validated assessment instrument."[29]

Other states have taken a more permissive approach to the use of recidivism risk assessment tools, simply encouraging or permitting consideration of risk assessments without explicitly requiring it.[30] For example, Louisiana law gives courts the option of using a "single presentence investigation validated risk and needs assessment tool prior to sentencing an adult offender eligible for assessment."[31] Washington State's sentencing law takes a similar approach, stating that the judge at a sentencing hearing "may order the department to complete a risk assessment report," except in certain cases.[32]

B. *Modern Risk Assessment Tools*

The increased use of risk assessment tools in sentencing is closely tied to the development of new modeling techniques utilizing machine learning,[33] the process by which a computer program is given a large quantity of data and tasked with identifying variables in the data that correlate with a specified outcome.[34] The use of machine learning has greatly increased in recent years, largely due to "the accumulation of large datasets for analysis and advances in computing power and machine learning theory that have enabled much more complex analysis of those datasets."[35] This technology is now used in myriad ways, including

---

28. Ky. Rev. Stat. Ann. § 532.007(3)(a)–(b) (West 2019).

29. Tenn. Code Ann. § 41-1-412(b) (2019). While several statutes include a requirement that the risk assessment tool be "validated," perhaps allaying concern that the tool will be inaccurate or biased, some jurisdictions have relied on outdated or unsubstantiated validation studies to meet this vague statutory requirement. See Derek Thompson, Should We Be Afraid of AI in the Criminal-Justice System?, Atlantic (June 20, 2019), https://www.theatlantic.com/ideas/archive/2019/06/should-we-be-afraid-of-ai-in-the-criminal-justice-system/592084/ [https://perma.cc/9WWZ-ZT6R].

30. See Kehl et al., supra note 15, at 16; see also State v. Loomis, 881 N.W.2d 749, 759 (Wis. 2016) (describing the recent rise of statutory risk assessment requirements in sentencing).

31. La. Code Crim. Proc. Ann. art. 15:326 (2018).

32. Wash. Rev. Code § 9.94A.500 (2019).

33. See, e.g., Kehl et al., supra note 15, at 9; Robert Brauneis & Ellen P. Goodman, Algorithmic Transparency for the Smart City, 20 Yale J.L. & Tech. 103, 113 (2018); Aziz Z. Huq, Racial Equity in Algorithmic Criminal Justice, 68 Duke L.J. 1043, 1045 (2019).

34. See David Lehr & Paul Ohm, Playing with the Data: What Legal Scholars Should Learn About Machine Learning, 51 U.C. Davis L. Rev. 653, 671 (2017) (defining machine learning as "an automated process of discovering correlations . . . between variables in a dataset, often to make predictions or estimates of some outcome").

35. Brauneis & Goodman, supra note 33, at 113.

in self-driving cars,[36] disease detection,[37] online-shopping product recom-
mendations,[38] and, of course, recidivism risk assessment tools.[39]

 The machine learning process proceeds through eight (sometimes
overlapping or repeated) steps, as identified by Professor Paul Ohm and
David Lehr and summarized in the table below.[40]

TABLE 1: THE MACHINE LEARNING PROCESS

Problem Definition	The developer determines the outcome the final algorithm should predict.[41] In the case of a recidivism risk assessment, this would be the likelihood of a given defendant committing another crime in the future.
Data Collection	The developer assembles a sufficiently large dataset from which the machine learning model can identify patterns.[42] For recidivism risk, this dataset would likely include extensive criminal histories collected from local government offices.
Data Cleaning	The developer combs through the dataset to identify incorrect or missing information.[43] Missing or incorrect values may require entire entries to be eliminated from the dataset, or replacement values to be imputed from the existing data. Although a single error may not be significant, it is impossible to assess the aggregate impacts of cleaning decisions on the overall dataset without detailed explanations from the developer.[44]

---

36. Russell Brandom, Self-Driving Cars Are Headed Toward an AI Roadblock, The
Verge (Jul. 3, 2018), https://www.theverge.com/2018/7/3/17530232/self-driving-ai-winter-
full-autonomy-waymo-tesla-uber [https://perma.cc/2G36-HR2B] (explaining the obstacles
that machine learning must overcome before self-driving cars can operate in a fully
autonomous capacity).

37. E.g., Ryan Poplin, Avinash V. Varadarajan, Katy Blumer, Yun Liu, Michael V.
McConnell, Greg S. Corrado, Lily Peng & Dale R. Webster, Prediction of Cardiovascular
Risk Factors from Retinal Fundus Photographs via Deep Learning, 2 Nature Biomedical
Engineering 158 (2018) (describing the use of machine learning software to assess risk of
cardiovascular disease through retinal image processing).

38. Frank Catalano, What's Not to 'Like?' Amazon Tests Machine-Learning Driven
Scout Instant Recommendation Engine, GeekWire (Sept. 19, 2018), https://www.geekwire.com
/2018/whats-not-like-amazon-tests-machine-learning-driven-scout-instant-recommendation-
engine/ [https://perma.cc/3MV2-DHKX] (detailing Amazon's recent implementation of the
Scout product-recommendation algorithm in the Amazon App).

39. Turgut Ozkan, Predicting Recidivism Through Machine Learning (May 2017)
(unpublished Ph.D. dissertation, University of Texas at Dallas) (manuscript at 63–66) (on
file with the *Columbia Law Review*) (comparing several machine learning models for their
efficacy in predicting recidivism).

40. See Lehr & Ohm, supra note 34, at 672–702.

41. See id. at 672–77.

42. See id. at 677–81.

43. Id. at 681–83.

44. Eaglin, supra note 17, at 80.

Summary Statistics Review	A developer conducts a review of the aggregate statistics to identify outliers.[45]
Data Partitioning	The dataset is divided into two subsets: one that will be used to "train" the algorithm and a second that will later be used to test its predictive ability.[46]
Model Selection	The developer chooses the model that will generate the predictive algorithm.[47] While all machine learning models work toward the same end—producing an algorithm with maximal predictive accuracy—there are several different methods for achieving this goal. Different models are better equipped to digest and predict quantitative or qualitative values, to over- or underestimate in their predictions, and to provide some amount of explanation for how results are calculated.[48]
Model Training	The selected model is applied to the subset of data that has been designated for training.[49] This is the "learning" portion of the process, when the model identifies patterns in the dataset and develops the predictive algorithm based on these patterns.[50] This process may be repeated several times, with developers fine-tuning both the data and the model over the course of multiple learning cycles. It is important to keep in mind that the ultimate objective of a machine learning model is not to identify inherent causal relationships in the dataset, but "to make classifications based on mathematical descriptions . . . that yield the lowest error rates."[51]
Model Deployment	The patterns that have been identified in the data are converted into a usable interface.[52] In the risk assessment context, this includes translating the "quantitative outcome into a qualitative 'risk score' used by criminal justice actors at sentencing."[53]

Today's risk assessment algorithms differ from earlier tools in their complexity and sophistication, and consequently, in their opacity.[54] This can make it difficult to understand the many decisions that go into

---

45. See Lehr & Ohm, supra note 34, at 683–84.

46. See id. at 684–88.

47. Id. at 688.

48. See id. at 688–95.

49. See id. at 695–701.

50. See Coglianese & Lehr, Regulating by Robot, supra note 11, at 1157 ("[T]hese algorithms make repeated passes through data sets, progressively modifying or averaging their predictions to optimize specified criteria.").

51. Id. at 1158.

52. See Lehr & Ohm, supra note 34, at 701–02.

53. Eaglin, supra note 17, at 85.

54. See infra section II.A.

creating them, particularly when these decisions are outsourced to private contractors.[55] As Professor Jessica Eaglin has observed, "Decisions about the data to collect, the recidivism event to observe, and the risk factors selected have great import to understanding what and how a resulting recidivism risk tool predicts."[56] For example, in the problem definition stage, if the algorithm developer chooses to define recidivism as an arrest occurring within five years of release, the algorithm will identify a different set of offenders in the dataset as recidivists than if the developer had selected a ten-year timeframe. Because this decision changes which offenders are classified as recidivists—and which are not—it inevitably influences the patterns that the model identifies in drawing connections between those offenders.[57]

C. *Constitutional Challenges to Algorithmic Risk Assessment*

The increasing use of risk assessment algorithms in criminal justice has been accompanied by vocal concern about the constitutionality of their operation and deployment. In a notable Wisconsin case from 2016, *State v. Loomis*,[58] Eric Loomis brought a due process challenge against the use of COMPAS, the most widely used risk assessment tool in America,[59] after he received a high risk score and was sentenced to six years in prison.[60] The Wisconsin Supreme Court held that the use of COMPAS did not violate due process, in part because Loomis had the opportunity to review the accuracy of the information that produced his risk score.[61] Despite the outcome in *Loomis*, the due process implications of government use of predictive algorithms in criminal adjudication remain unsettled.[62] And while Loomis did not explicitly bring an equal protection

---

55. See supra notes 7–9 and accompanying text.

56. Eaglin, supra note 17, at 73.

57. For a more in-depth discussion of how the decisions of private developers shape the predictions made by risk assessment algorithms, see infra section II.B.

58. 881 N.W.2d 749 (Wis. 2016).

59. Raso et al., supra note 20, at 23.

60. See *Loomis*, 881 N.W.2d at 756 n.18. For further discussion of *Loomis*, see infra section II.A; see also Recent Case, supra note 22. For an extended exploration of the due process implications of automated decisionmaking, see generally Danielle Keats Citron & Frank Pasquale, The Scored Society: Due Process for Automated Predictions, 89 Wash. L. Rev. 1 (2014) (elaborating on the need for procedural regularity in the use of automated predictions); Danielle Keats Citron, Technological Due Process, 85 Wash. U. L. Rev. 1249 (2008) (describing the lack of procedural safeguards in automated decisionmaking systems).

61. See *Loomis*, 881 N.W.2d at 761.

62. See Garrett & Monahan, supra note 21, at 2 ("The implications of this due process analysis for settings in which judges are informed by quantitative risk assessment methods have not been fully explored."); see also Loomis v. Wisconsin, 137 S. Ct. 2290 (2017) (mem.) (denying certiorari in *Loomis*).

claim,[63] a number of scholars have identified potential equal protection challenges arising from the use of risk assessment algorithms, which may consider protected characteristics in determining risk scores and are known to have disparate impacts on certain groups.[64]

As the following Part demonstrates, the private development of risk assessment algorithms inhibits investigation of the merits of these constitutional claims. Even when public actors apply privately developed risk assessment tools in sentencing determinations, issues of transparency and accountability persist.[65] Judges struggle to understand and interrogate algorithmic results, allowing private actors to influence sentencing outcomes without being subject to traditional accountability mechanisms.[66] After all, it is difficult, if not impossible, to enforce constitutional guarantees when the violation is obscured from even the judge's view.[67]

---

63. See *Loomis*, 881 N.W.2d at 766 ("Loomis does not bring an equal protection challenge in this case. Thus, we address whether Loomis's constitutional due process right not to be sentenced on the basis of gender is violated if a circuit court considers a COMPAS risk assessment at sentencing.").

64. See, e.g., Kehl et al., supra note 15, at 23–26; Melissa Hamilton, Risk-Needs Assessment: Constitutional and Ethical Challenges, 52 Am. Crim. L. Rev. 231, 242–63 (2015); Huq, supra note 33, at 1083–102; Anne L. Washington, How to Argue with an Algorithm: Lessons from the COMPAS-ProPublica Debate, 17 Colo. Tech. L.J. 131, 150–51 (2018). Interestingly, at least one state has explicitly required the use of gender-specific risk assessment methodology at sentencing, adding another wrinkle to the equal protection issues raised by the use of this technology. See Conn. Gen. Stat. § 18-81z (2019) (requiring the Connecticut Department of Correction to develop a risk assessment tool that will "incorporate use of both static and dynamic factors and utilize a gender-responsive approach that recognizes the unique risks and needs of female offenders").

65. Cf. Nina A. Mendelson, Private Control over Access to Public Law: The Perplexing Federal Regulatory Use of Private Standards, 112 Mich. L. Rev. 737, 787–90 (2014) [hereinafter Mendelson, Private Control] ("[O]bstacles to access by the public and by Congress to the rules' content will obviously impair a congressional committee's ability to perform any meaningful oversight . . . [and] undermine[] the usability of accountability mechanisms for ordinary people."). As Mendelson explains, the Code of Federal Regulations "contains nearly 9,500 'incorporations by reference' of standards" developed primarily by private organizations. Id. at 739. Through the process of incorporation by reference, privately developed standards become public law, but are not fully incorporated into the text of the C.F.R. Instead, the incorporating agency typically "refers the reading public to the [standards development organization,] . . . [which] seemingly without exception, assert[s] copyright protection and an entitlement to charge a 'purchase price' for access. . . . [T]he reader's alternative is to make an appointment at the OFR's reading room in Washington, D.C. The reading room contains no photocopier." Id. at 743. For more information on the invocation of copyright protections by standards development organizations, see Peter L. Strauss, Private Standards Organizations and Public Law, 22 Wm. & Mary Bill Rts. J. 497, 499–518 (2013).

66. See infra sections II.A–.B.

67. See infra notes 74–77 and accompanying text.

II. PRIVATIZATION BY AUTOMATION

As seen in Part I, a number of constitutional concerns arise from the use of recidivism risk assessment algorithms in sentencing.[68] In the outpouring of scholarship on the constitutional problems presented by these algorithms, however, one question has been largely overlooked: Who is accountable for ensuring the constitutional compliance of risk assessment systems? This Part outlines the gap in constitutional accountability that arises from the use of privately developed risk assessment algorithms in sentencing, focusing on the legal and technological obstacles judges face in applying algorithmic risk scores.

Section II.A begins with a summary of these obstacles, looking at the opacity and false sense of objectivity that surround algorithmic decisionmaking and potentially lead judges to rely heavily on the results of risk assessment tools. While previous scholarship has discussed these opacity and objectivity problems in and of themselves, section II.B demonstrates how these obstacles obscure the role that private developers play in shaping a risk assessment algorithm, giving them undue influence in sentencing determinations. Section II.C concludes this Part by identifying the gap in constitutional accountability that arises when the decisions of risk assessment developers are not clear to the judges tasked with applying their tools.

A.   *Limits on Judicial Understanding of Risk Scores*

This section summarizes the difficulties judges face in understanding and applying risk assessment algorithms.[69] Both legal and technological obstacles may obscure essential information about the development of these tools and the meaning of their outputs. Due in part to this opacity, judges may be inclined to rely on algorithmic predictions, which appear scientific and objective when little background information is available.

1. *Legal and Technological Opacity.* — Two kinds of opacity limit judges' ability to understand the operation of risk assessment algorithms: legal opacity and technological opacity. Legal opacity refers to legal obstacles, such as trade secret protections,[70] that prevent a judge from

---

68. See supra section I.C.

69. Northpointe, the producer of the widely used COMPAS risk assessment algorithm at issue in *State v. Loomis*, has openly acknowledged the difficulties of interpreting and applying COMPAS results. See Northpointe, Practitioner's Guide to COMPAS Core 4 (2015), http://www.northpointeinc.com/downloads/compas/Practitioners-Guide-COMPAS-Core-_031915.pdf [https://perma.cc/Q23R-7849] ("Interpretation [of risk assessments] is a skill that needs to be honed over time.").

70. Courts have only recently begun to recognize trade secret protections in criminal cases. See Rebecca Wexler, Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System, 70 Stan. L. Rev. 1343, 1388–95 (2018). Although in civil cases these protections can be overcome through protective orders or in camera review, the use

accessing information about the algorithm,[71] whereas technological opacity relates to a judge's inability to understand available information about the algorithm due to a lack of relevant expertise.[72] This section argues that, without transparency in the development and function of risk assessment algorithms, judges are unable to understand and properly apply their results—an idea best demonstrated by *State v. Loomis*, the Wisconsin case discussed in Part I.[73]

In *Loomis*, the defendant challenged the use of the privately developed COMPAS risk assessment tool in determining his six-year prison sentence.[74] On appeal to the Wisconsin Supreme Court, Loomis argued that the sentencing court's decision incorporated information that was unavailable to him because the creators of COMPAS, invoking trade secret protections, were able to avoid disclosing how the algorithm determined his risk score.[75] In upholding the sentence, the court stated that, unlike the due process violation that occurs when a sentencing court relies on information that the defendant is barred from reviewing, Loomis had access to the same material as the court that sentenced him.[76] Essentially, no due process issue existed because the sentencing judge and the defendant had equally limited information about the operation of the algorithm.[77]

---

of these techniques in the criminal context may conflict with a defendant's Sixth Amendment right to a public trial. See id. at 1353 n.46.

71. See infra notes 73–79 and accompanying text; see also Deirdre K. Mulligan & Kenneth A. Bamberger, Saving Governance-by-Design, 106 Calif. L. Rev. 697, 720 (2018) ("[C]losed-source code leaves outsiders 'unable to discern how a system operates and protects itself' and shields unintended errors that distort even clear legal and managerial goals." (footnote omitted) (quoting Danielle Keats Citron, Open Code Governance, 2008 U. Chi. Legal F. 355, 357)).

72. See infra notes 80–88 and accompanying text.

73. See supra notes 58–63 and accompanying text.

74. See State v. Loomis, 881 N.W.2d 749, 756 & n.18 (Wis. 2016). Because many details of how COMPAS was developed and operates are not public, it is unclear whether this tool was developed using machine learning processes. See Coglianese & Lehr, Regulating by Robot, supra note 11, at 1205 n.232. The problems arising from its use are nevertheless illustrative of the challenges judges face in applying these tools.

75. See *Loomis*, 881 N.W.2d at 761 ("Northpointe . . . considers COMPAS a proprietary instrument and a trade secret. Accordingly, it does not disclose how the risk scores are determined . . . . Loomis asserts that because COMPAS does not disclose this information, he has been denied information which the circuit court considered at sentencing.").

76. See id. Justice Shirley Abrahamson, in her concurrence, raised separate concerns about the lack of available information about the risk assessment algorithm at issue in the case. See id. at 774 (Abrahamson, J., concurring) ("[T]his court's lack of understanding of COMPAS was a significant problem in the instant case. At oral argument, the court repeatedly questioned both the State's and defendant's counsel about how COMPAS works. Few answers were available.").

77. Wexler, supra note 70, at 1346 ("The court reasoned that no due process violation had occurred in part because the judge's own access to the secrets was equally limited.").

Much of the scholarship discussing *Loomis* has focused on the successful invocation of trade secret privileges by Northpointe,[78] the developer of COMPAS, to prevent disclosure of how the algorithm operates.[79] Although trade secret protections were a central issue in *Loomis*,[80] the failure to disclose proprietary source code is only one of many ways in which the process for developing these tools has been obscured. In addition to the *legal* opacity arising from the use of trade secret protections to prevent source code disclosure, there is a layer of *technological* opacity working to obscure the ways in which risk assessment algorithms make their determinations.

While the trade secret obstacle certainly limits transparency, the opacity arising when judges with little to no technical knowledge apply incredibly complex software is equally troublesome. Even if trade secret protections were eliminated and the source code for these tools were provided to both judges and defendants, it is unlikely that this disclosure would meaningfully increase their understanding of how these tools function.[81] Whereas the significance and weight of factors used in traditional statistical models can be clearly explained, there is currently no practical way to demonstrate how a given input—such as the offender's age—influences the risk score determined by an algorithm.[82] Looking back at the reasoning in *Loomis*,[83] it is easy to imagine a situation in which the court and defendant receive source code that they have no means of interpreting, but the judge finds no due process issue because there is equal access to information for all involved. It is therefore important to ensure that judges understand the inner workings of the technology that they apply, especially when, as in cases like *Loomis*, that technology may play a central role in sentencing.

---

78. Following a merger in 2017, the company that owns and licenses COMPAS is now known as Equivant. See Equivant FAQs, Equivant, https://www.equivant.com/faq/ (on file with the *Columbia Law Review*) (last visited July 24, 2019). However, because both the *Loomis* case and subsequent scholarship have largely referred to the company as Northpointe, this Note does so to avoid confusion.

79. See Natalie Ram, Innovating Criminal Justice, 112 Nw. U. L. Rev. 659, 683–86 (2018) ("[A]lthough Northpointe has disclosed the 137-question survey that provides informational input for its program, it has refused to disclose how that information is used or weighted to arrive at a particular recidivism risk score."); see also Wexler, supra note 70, at 1368–71; Carlson, supra note 14, at 315–29.

80. See *Loomis*, 881 N.W.2d at 761.

81. See Coglianese & Lehr, Regulating by Robot, supra note 11, at 1190 ("Most potential counsel or agency hearing examiners do not possess the necessary skills to interrogate machine learning systems.").

82. See id. at 1159–60.

83. See supra notes 74–77 and accompanying text.

A large source of the technological opacity described above is what has been called the "black box" nature of machine learning[84]—the idea that machine learning algorithms transform a set of inputs, such as criminal history, into an output, like a recidivism risk score, through a process that is not easily explained.[85] When the algorithm used to assess a defendant's recidivism risk is created through machine learning, rather than by a human statistician, it is not always clear which of the data points provided as inputs actually factor into the final output, or how frequently and heavily those inputs are considered in the algorithm's complex calculus.[86] Consequently, when a judge or defendant receives an offender's algorithmically generated risk score, they "may not understand why someone is considered to have a low, medium, or high risk of recidivism."[87] Because "[t]he results of machine learning analysis are not intuitively explainable," they "cannot support causal explanations of the kind that underlie the reasons traditionally offered to justify governmental action."[88]

2. *The False Objectivity of Data.* — A second problem arising from the use of risk assessment algorithms in sentencing is the false sense of objectivity that surrounds algorithmic predictions.[89] Professor Kate Crawford has described this phenomenon as "data fundamentalism," or "the notion that correlation always indicates causation, and that massive data sets and predictive analytics always reflect objective truth."[90] Because of the pervasive idea that all data is scientific, and that processes like

---

84. See Coglianese & Lehr, Regulating by Robot, supra note 11, at 1159 ("[M]achine-learning algorithms are often described as transforming inputs to outputs through a black box."). See generally Frank Pasquale, The Black Box Society: The Secret Algorithms that Control Money and Information (2015) (arguing for greater transparency, accountability, and regulation in the pervasive use of personal data).

85. See Eaglin, supra note 17, at 106 ("Unlike humans, the tools provide no explanation for their results other than the numerical outcomes translated into risk scores."); see also Coglianese & Lehr, Regulating by Robot, supra note 11, at 1159–60 ("An analyst cannot look inside the black box to understand how that transformation occurs or describe the relationships with the same intuitive and causal language often applied to traditional statistical modeling.").

86. See Eaglin, supra note 17, at 119 ("Risk tools using this modeling create difficult interpretability issues, as the developers creating the tools cannot explain what factors a tool uses to predict recidivism risk."); see also Coglianese & Lehr, Regulating by Robot, supra note 11, at 1167 ("The results of algorithms do not depend on humans specifying in advance how each variable is to be factored into the predictions . . . . These algorithms effectively look for patterns on their own.").

87. Eaglin, supra note 17, at 110.

88. Coglianese & Lehr, Regulating by Robot, supra note 11, at 1167.

89. See Jasmine Sun, I Won't Pledge Allegiance to Big Data, Stan. Daily (Jan. 15, 2018), https://www.stanforddaily.com/2018/01/15/i-wont-pledge-allegiance-to-big-data/ [https://perma.cc/7SE3-X8PH] ("[T]he rise of data-driven decision-making has been accompanied by a dangerous ruse of objectivity . . . .").

90. Kate Crawford, The Hidden Biases in Big Data, Harv. Bus. Rev. (Apr. 1, 2013), https://hbr.org/2013/04/the-hidden-biases-in-big-data [https://perma.cc/RY5E-FTV3].

machine learning are inherently logical, the unique attributes of modern risk assessment tools "combine to make machine-learning techniques appear qualitatively more independent from humans when compared to other statistical techniques."[91] When interpreting the output of a machine learning algorithm, it can be tempting to view the results as objective, fair, and inevitable simply because they were produced by a computer.[92]

The guise of objectivity is an issue whenever data is used to supplement human decisionmaking, but it is particularly concerning when judges use proprietary algorithms to inform their judgments. Judges, directed by statute to consider the results of risk assessment algorithms in their sentencing decisions,[93] might place strong emphasis on risk scores.[94] As Professors Robert Brauneis and Ellen P. Goodman explain, judges "are expected to exercise human judgment over algorithmic predictions so that they may catch faulty predictions. In theory, the algorithmic edict is advisory only. In practice, decisionmakers place heavy reliance on the numbers, raising the stakes for their fairness."[95] Judicial reliance on algorithmic predictions is especially problematic when there has been little to no independent validation of their results.[96]

This false sense of objectivity is connected to the opacity problems discussed above, which also contribute to judicial reliance on algorithmic risk scores. As Eaglin explains, private risk assessment developers are

---

91. Coglianese & Lehr, Regulating by Robot, supra note 11, at 1167; see also Raso et al., supra note 20, at 22 ("[T]he objective veneer that coats the outputs of these tools obscures the subjective determinations that are baked into them.").

92. See Anupam Chander, The Racist Algorithm?, 115 Mich. L. Rev. 1023, 1034 (2017) (reviewing Pasquale, supra note 84) ("[W]hen algorithms replace human decisionmaking, algorithms give the decisionmaking 'a patina of inevitability' . . . . Algorithms can make decisionmaking seem fair precisely because computers are logical entities which should not be infected by all-too-human bias. But that would be an unwarranted assumption . . . ." (citation omitted) (quoting Pasquale, supra note 84)); see also Andrea Roth, Machine Testimony, 126 Yale L.J. 1972, 1992 (2017) ("Human design choices—unless disclosed to the factfinder—can lead to inferential error if a machine's conveyance reflects a programmed tolerance for uncertainty that does not match the one assumed by the factfinder."); Jones, supra note 19 (explaining that if a risk assessment tool "simply repackages racial disparity into a seemingly objective score, we are worse off than when we started").

93. See supra notes 24–32 and accompanying text.

94. See Collins, supra note 21, at 68 ("[J]udges who receive predictive risk information may modify their sentence in the direction of the risk prediction. A few empirical studies support this inference." (footnote omitted)).

95. Brauneis & Goodman, supra note 33, at 123. As the authors note, "When the 'machine says so,' it can be difficult for rushed and over-extended human decision makers to resist the edict." Id. at 126. Other scholars have suggested that sentencing judges may rely heavily on risk assessment tools to avoid political backlash that might arise if an offender labeled "high risk" by the algorithm is released and then commits another crime. See Kehl et al., supra note 15, at 14.

96. See supra note 29.

incentivized to avoid disclosing details related to their tools, since this information may undermine the credibility of their software, and therefore their competitiveness.[97] It is possible that, rather than relying on the results of an opaque machine because it seems "scientific," judges may disregard a risk score when its origins are unclear.[98] However, many scholars have suggested that, despite the lack of transparency, decisionmakers do in fact rely on the results of risk assessment tools in making sentencing determinations.[99]

Although these opacity and objectivity problems make sense in theory, it is worthwhile to examine the reality of how judges use and apply risk scores. Another case from Wisconsin provides a useful example. Paul Zilly agreed to a plea deal with prosecutors to serve one year in jail for theft.[100] After seeing his risk assessment score, a judge rejected the deal and sentenced Zilly to two years in prison along with three years of supervision, noting that his risk score was "about as bad as it could be."[101] Although the judge later reduced Zilly's sentence, he

---

97. See Eaglin, supra note 17, at 111–12 ("Competition amongst tool creators . . . encourages developers to remain vague about the subjective judgments embedded in their tools. Disclosing specific information about tool-construction choices may lead a consumer to perceive the underlying data set as methodologically weak or unsound, and ultimately seek out another product."); see also Carlson, supra note 14, at 315 ("As a result of the growing trend to implement actuarial risk assessment in sentencing, risk assessment has become a competitive industry . . . .").

98. This raises an interesting question: To what extent should judges rely on these tools in making their determinations? If a judge always deviates from the algorithm's recommendation, that is presumably inconsistent with the statutory purpose of requiring the risk assessment. However, if a judge's determination always aligns with that of the algorithm, that could indicate that the algorithm has completely supplanted the judge's decisionmaking, or that the algorithm fully aligns with the judge's intuition and therefore serves no real purpose. See Brauneis & Goodman, supra note 33, at 127 ("If the algorithm is opaque, the government official cannot know how to integrate its reasoning with her own, and must either disregard it, or follow it blindly."). Although outside the scope of this Note, determining what amount of deviation, if any, from an algorithm's recommendation is appropriate is a rich question for further exploration.

99. See, e.g., Harry Surden, Values Embedded in Legal Artificial Intelligence (Univ. of Colo. Law, Legal Studies Research Paper No. 17-17, 2017), https://ssrn.com/abstract_id=2932333 (on file with the *Columbia Law Review*) (explaining that decisionmakers may "give more deference to computer-based recommendations, in contrast to comparable human-based assessments, given the aura of mechanistic objectivity surrounding computer-generated, analytics-based, analyses"); see also Collins, supra note 21, at 68 & n.64. Some have even suggested that the appearance (as opposed to the *existence*) of objectivity in algorithmic decisionmaking may motivate lawmakers to adopt these tools in order to shield subjective decisions from scrutiny. See Jeremy Isard, Note, Under the Cloak of Brain Science: Risk Assessments, Parole, and the Powerful Guise of Objectivity, 105 Calif. L. Rev. 1223, 1252 (2017) ("By . . . rendering scientific inquiry and agency findings one and the same, [the Board of Parole Hearings] has effectively rendered the components of its suitability inquiry mutually affirming and impervious to legal challenge.").

100. See Angwin et al., supra note 20.

101. Id.

explained that, "[h]ad I not had the COMPAS, . . . it would likely be that I would have given one year, six months."[102] While this evidence is admittedly anecdotal, it suggests that judges may rely heavily on risk scores to inform their sentencing decisions, imposing harsher sentences on those determined to present a high risk of recidivism. Scholars and journalists have noted this issue, raising alarms about allowing machines to determine prison sentences,[103] but the problem is more complicated than that. A focus on the algorithm as decisionmaker distracts from the human actors that have come to play a major role in sentencing determinations: the private developers of risk assessment tools.

B. *The Privatization Problem*

The opacity and objectivity problems discussed above prevent judges from understanding the inner workings of risk assessment tools, thereby obscuring the subjective decisions that shape these algorithms and increasing the likelihood that those decisions will influence someone's sentence.[104] This means that private developers play a significant part in sentencing determinations without being subject to traditional constitutional accountability mechanisms.[105] As Brauneis and Goodman explain, in providing algorithmic tools to government bodies, "private entities assume a significant role in public administration. [Critical information] comes to reside in the impenetrable brains of private vendors while the government, which alone is accountable to the public, is hollowed out, dumb and dark."[106]

---

102. Id.

103. See, e.g., Ellora Thadaney Israni, Opinion, When an Algorithm Helps Send You to Prison, N.Y. Times (Oct. 26, 2017), https://www.nytimes.com/2017/10/26/opinion/algorithm-compas-sentencing-bias.html (on file with the *Columbia Law Review*) ("But shifting the sentencing responsibility to a computer does not necessarily eliminate bias; it delegates and often compounds it."); Maayan Perel, Technological Reliefs: The Devil Is in the Technological Details 4 & nn.30–31 (unpublished manuscript) (on file with the *Columbia Law Review*) ("Delegation of *judicial* decision-making to proprietary algorithms, however, has been hardly explored. Mostly discussed in this context is the limited use of big data and machine-learning algorithms in criminal cases for risk assessment purposes." (footnote omitted)).

104. See Brauneis & Goodman, supra note 33, at 122 ("Unless the algorithmic prediction is self-executing, human beings have to understand the prediction in order to choose how much weight to give it in the decision-making process.").

105. Cf. Coglianese & Lehr, Regulating by Robot, supra note 11, at 1177 ("An algorithm, by its very definition, must have its parameters and uses specified by humans, and this property will likely prove pivotal in the legal assessment of specific applications of artificial intelligence by federal administrative agencies.").

106. Brauneis & Goodman, supra note 33, at 109. Recent privatization scholarship argues that government is not in fact "hollowed out" but rather aggrandized by privatization, which enables the outsourcing branch to exercise greater discretion than it would otherwise be able to. See Jon D. Michaels, Constitutional Coup: Privatization's Threat to the American Republic 119–41 (2017) [hereinafter Michaels, Constitutional Coup]. However, this argument may not extend to the case of privately developed risk

Returning to the eight steps of machine learning described by Lehr and Ohm,[107] it is clear that risk assessment developers make policy decisions throughout the machine learning process that can substantially impact the results of their algorithm.[108] For example, developers must quantitatively define what "recidivism" means in order for the model to be able to predict it, a process which can implicate major policy questions in criminal law.[109] A developer could instruct the model to predict recidivism based on whether someone will be arrested or convicted, or to predict the likelihood of this event occurring within one, five, or ten years.[110] While these subjective, private determinations have obvious significance for the meaning of a defendant's risk score,[111] they are not at all clear to a judge applying the risk assessment tool.

The collection and cleaning of the dataset from which the model learns necessarily present further opportunities for developers to influence the tool's predictive outcomes. After all, "[n]o predictive tool is better than the data set from which it originates."[112] Private developers make decisions about how to gather the data that will form the foundation of the risk assessment tool—including the jurisdictions to pull data from and the size of the dataset—which can have significant impacts on how the algorithm functions.[113] Determinations about which

---

assessment tools, in which government actors provide contractors with minimal policy guidance and lack the ability to scrutinize the actions of contractors for compliance. See infra section IV.A. Whereas the privatization Michaels describes allows government to use outsourcing to operate outside of traditional constraints, to sidestep noncompliant bureaucrats in favor of more willing contractors, or to entrench policy choices through long-term contracts, Jon D. Michaels, Privatization's Pretensions, 77 U. Chi. L. Rev. 717, 719–21 (2010), there is no indication that the move to privately developed risk assessment algorithms is an attempt to do these things. Rather, the use of privately developed risk assessment algorithms appears to be one of the "exceptions" to Michaels's concept of privatization, in which the contractor's expertise "far outstrips the government's," aligning more closely with the classic view of privatization as abdication. See Michaels, Constitutional Coup, supra, at 124.

107. See supra Table 1.

108. Eaglin, supra note 17, at 64 ("[A]ctuarial risk tools . . . reflect normative judgments familiar to sentencing law and policy debates. Yet . . . it is difficult to identify the normative judgments reflected in the information produced by the tools.").

109. See id. at 75 ("Framing this question requires that developers understand the objectives and requirements of the problem and convert this knowledge into a data problem definition[,] . . . requiring developers to finesse a social dilemma such that a computer can automate a responsive answer.").

110. See id. at 75–78.

111. Cf. Simon Chandler, Big Data Can't Bring Objectivity to a Subjective World, TechCrunch (Nov. 18, 2016), https://techcrunch.com/2016/11/18/big-data-cant-bring-objectivity-to-a-subjective-world/ [https://perma.cc/3F3S-VGSN] (arguing that while the validity of a predictive tool may be tied to its underlying data, "the deeper issue is the inevitable variation in how people classify these words themselves").

112. Eaglin, supra note 17, at 72.

113. See id. (explaining that decisions made while assembling the training dataset form the basis for a predictive algorithm and "have a significant impact on the outcomes

data points to exclude and how to replace erroneous or missing values all reflect subjective judgments about the underlying data.[114]

The process of assembling the risk assessment algorithm presents another set of opportunities to influence a tool's output. Not every predictive factor that a model identifies will be incorporated into the tool that is ultimately packaged and sold to the government. Private developers must therefore judge which factors to include and which to omit,[115] decisions which are often made without taking relevant state sentencing laws into consideration.[116] Furthermore, an algorithm's risk determination must ultimately be converted into a digestible output for judges to use, which requires developers to make additional policy judgments about what constitutes low, medium, and high risk.[117]

Because of both legal and technological opacity, as well as the veil of objectivity that surrounds algorithmic decisionmaking, most of the actions of these private developers are obscured from public view.[118] Although private actors influence the results of risk assessment algorithms at every stage of their development, these choices are not immediately clear to the judges applying these tools. As Eaglin notes, "[P]revious efforts to estimate risk at sentencing—like guidelines and mandatory penalties—made normative judgments about how to sentence and why, [but] those choices were apparent on the face of the mechanized tool. . . . With actuarial risk tools, normative judgments are more difficult or even impossible to discern."[119]

---

of these tools"). This precise issue was raised at trial in *Loomis,* when Dr. David Thompson, an expert witness for the defendant, testified that "[t]he Court does not know how the COMPAS compares that individual's history with the population that it's comparing them with. The Court doesn't even know whether that population is a Wisconsin population, a New York population, a California population." State v. Loomis, 881 N.W.2d 749, 756 (Wis. 2016). Dr. Thompson went on to testify that "[t]here's all kinds of information that the court doesn't have, and what we're doing is we're mis-informing the court when we put these graphs in front of them and let them use it for sentenc[ing]." Id. at 756–57.

114. See Eaglin, supra note 17, at 80 ("Because data sets originate from a variety of sources, information provided may be incorrect. . . . Researchers seeking to use that information will either 'fix' the information or throw the defendant out of the data set. 'Fixing' the information requires subjective judgments about what the information likely means." (footnote omitted)).

115. Id. at 83.

116. Id. ("Tool creators tend to include predictive factors without reference to whether their use is regulated in state sentencing systems.").

117. See id. at 87 ("This decision requires some expertise not only in what the tool is predicting, but also in how society interprets the numerical outcome's meaning. In short, where developers place cut-off points reflects a normative judgment about how much likelihood of risk is acceptable in society without intervention." (footnotes omitted)); cf. Roth, supra note 92, at 1992 (discussing the ambiguity that arises when information about a DNA-matching program's error rate is not disclosed at trial).

118. Surden, supra note 99 (explaining that a lack of transparency "masks an underlying series of subjective judgments on the part of the system designers").

119. Eaglin, supra note 17, at 88.

C.   *The Constitutional Accountability Gap*

At the heart of this problem is the principle of constitutional accountability—the idea that "the Constitution imposes limits on the actions that governments can take . . . [, and that] individuals injured by exercises of government power can enforce these constitutional limits in court."[120] As Professor Paul R. Verkuil explains, "Accountability for acts of government is difficult when duties are delegated to private hands and secrecy covers the tracks."[121] Because the Constitution "erects no shield against merely private conduct, however discriminatory or wrongful,"[122] the privatization of government functions raises questions about how to ensure that constitutional accountability is preserved in the now-private performance of previously public functions.[123]

A central facet of constitutional accountability is the ability of individuals to enforce "constitutional restrictions in court through judicial review."[124] As the preceding sections demonstrate, the opacity surrounding privately developed risk assessment algorithms, along with the false sense of objectivity that these algorithms provide, combine to prevent judges from reviewing the decisions of private actors that are embedded in these technologies.[125] While judges have a symbolic ability to review and deviate from the determinations of risk assessment algorithms, the issues raised in section II.A effectively limit the extent to which this is possible.[126] Without additional information about how these algorithms

---

120. Gillian E. Metzger, Privatization as Delegation, 103 Colum. L. Rev. 1367, 1373 (2003).

121. Paul R. Verkuil, Outsourcing Sovereignty: Why Privatization of Government Functions Threatens Democracy and What We Can Do About It 13 (2007); see also Metzger, supra note 120, at 1400 ("[T]he move to greater government privatization poses a serious threat to the principle of constitutional accountability . . . [which] lies at the bedrock of U.S. constitutional law."); cf. Fuentes v. Shevin, 407 U.S. 67, 93 (1972) (explaining that when laws "abdicate effective state control over state power . . . [t]he State acts largely in the dark").

122. Shelley v. Kraemer, 334 U.S. 1, 13 (1948). The Thirteenth Amendment is a notable exception to this principle. See U.S. Const. amend. XIII (outlawing slavery and involuntary servitude in terms that extend to both government and private actors).

123. See Metzger, supra note 120, at 1401–03 (explaining that in instances when government functions are delegated to private actors there is concern that "handing over government programs to private entities will operate to place these programs outside the ambit of constitutional constraints, given the Constitution's inapplicability to 'private' actors"); see also Jody Freeman, The Private Role in Public Governance, 75 N.Y.U. L. Rev. 543, 574 (2000) ("As nonstate actors, [private actors] remain relatively insulated from the legislative, executive, and judicial oversight to which agencies must submit.").

124. Metzger, supra note 120, at 1401–02.

125. See supra sections II.A–.B.

126. Critics have raised concerns about the potential for due process violations to arise from "arbitrariness-by-algorithm" when these unexplainable technologies are used in consequential government decisionmaking. See, e.g., Citron & Pasquale, supra note 60, at 11–13, 24. For a more detailed, albeit dated, exploration of the potential for arbitrary decisionmaking in predictions of dangerousness in criminal adjudication, see Michael

are developed, it is impossible for judges to know whether they represent due process or equal protection violations,[127] or to use them in a way that is consistent with the intent of the legislature.[128]

The private action here is subtly obscured but raises the same issues as other instances of privatization of government functions[129]—namely that the mechanisms of government oversight in these relationships are structured in a way that does not comport with notions of constitutional accountability. Without the ability to more fully understand how these tools operate, judges are unable to interrogate their results in a way that sufficiently preserves constitutional protections.[130] As Lehr and Ohm explain, "Only one who is attentive to the many ways in which data can be selected and shaped—say, during data cleaning or model training— will characterize fully the source of the stink."[131]

The use of privately developed risk assessment algorithms in sentencing is, therefore, an instance in which "private market providers are cloaked in state clothes," giving rise to potentially harmful conduct in need of greater constitutional oversight.[132] Implementations of machine learning technology in government decisionmaking obscure private action behind a technological veil, masking the reality of the situation— that the developers of these tools have been given inordinate power in sentencing, power that should be subject to traditional constitutional limitations.[133] While source code disclosure may nominally give both defendants and judges the ability to understand how these algorithms operate, the chance that a judge or criminal defendant will be able to

---

Tonry, Prediction and Classification: Legal and Ethical Issues, *in* Prediction and Classification: Criminal Justice Decision Making 367, 377–81 (Don M. Gottfredson & Michael Tonry eds., 1987).

127. See supra section I.C.

128. See supra note 98 (explaining the difficulty of determining whether a judge's use of a risk assessment algorithm comports with legislative intent).

129. Brauneis & Goodman, supra note 33, at 119 ("The idea that algorithms are a science without politics can obscure the stakes of their private control that are clearer in other areas of privatization, such as schools and prisons.").

130. See Lehr & Ohm, supra note 34, at 656.

131. Id.; see also Brauneis & Goodman, supra note 33, at 109 ("When a government agent implements an algorithmic recommendation that she does not understand and cannot explain, the government has lost democratic accountability . . . .").

132. Sacha M. Coupet, The Subtlety of State Action in Privatized Child Welfare Services, 11 Chap. L. Rev. 85, 93 (2007) (discussing the need for increased constitutional accountability in public–private partnerships for the provision of child welfare services).

133. Scholars have recently noted that "[a]utomation is intensifying the privatization of the justice system," e.g., Wexler, supra note 70, at 1349, arguing that the growing government interest in advanced technology is increasing the role of private actors in the performance of government work. But the solutions proposed—such as a reworking of trade secret protections to eliminate obstacles to source code disclosure, see id. at 1413– 29—ultimately fail to address the ways in which government investment in automation alters existing structures of government accountability.

make sense of this technical information in order to mount or assess a constitutional challenge is extremely low.[134] Rather than solely focusing on the constitutionality of a risk assessment algorithm or the defendant's ability to challenge trade secret protections, it is equally important to examine the judge's ability to apply the risk assessment in a way that allows for enforcement of constitutional guarantees.[135]

III. THE PRIVATE DELEGATION DOCTRINE

Responding to the gap in constitutional accountability described in Part II, this Part turns to the private delegation doctrine as a framework for understanding the shortcomings of existing laws that govern the use of privately developed risk assessment algorithms in sentencing. Section III.A begins with a description of the traditional nondelegation doctrine, which limits delegations to public actors, and then discusses the justifications for applying similar, and possibly heightened, scrutiny when government power is placed in the hands of private parties. Section III.B summarizes the origins and evolution of the private delegation doctrine, which limits the ability of Congress to delegate authority to private actors. Section III.C then identifies key principles that guide courts' analyses of private delegations. As Part IV then explores, these private delegation principles provide a framework for understanding the problems presented by privately developed risk assessment algorithms, as well as a path to restoring accountability to their use.

A. *Private Versus Public Delegation*

In administrative law, the nondelegation doctrine limits Congress's ability to transfer power to administrative agencies and other actors. While the doctrine formally states that Congress may not delegate legislative power at all,[136] congressional grants of power are rarely invalidated in practice.[137] Recognizing that strict limitations on delegation are impracti-

---

134. See supra notes 80–82 and accompanying text.

135. See Coglianese & Lehr, Regulating by Robot, supra note 11, at 1154–55 ("[P]ublic officials, lawyers, and judges should ask how well the use of machine learning will conform to well-established legal principles of constitutional and administrative law. . . . [M]achine learning could be implemented irresponsibly in ways that, even though legal, might still offend more conventional notions of good government.").

136. Whitman v. Am. Trucking Ass'ns, 531 U.S. 457, 472 (2001) ("[T]he Constitution vests '[a]ll legislative Powers herein granted . . . in a Congress of the United States.' This text permits no delegation of those powers . . . ." (second alteration in original) (quoting U.S. Const. art. I, § 1)).

137. See id. at 474–75 ("[W]e have 'almost never felt qualified to second-guess Congress regarding the permissible degree of policy judgment that can be left to those executing or applying the law.'" (quoting Mistretta v. United States, 488 U.S. 361, 416 (1989) (Scalia, J., dissenting))); see also Keith E. Whittington & Jason Iuliano, The Myth of the Nondelegation Doctrine, 165 U. Pa. L. Rev. 379, 404 (2017) ("A review of the Court's [jurisprudence] . . . does not provide much basis for thinking that there was ever a

cal,[138] the modern doctrine requires that the legislature prescribe suffi-
cient policies and standards to restrict the scope of discretion that actors
have in wielding the power delegated to them.[139] When reviewing congres-
sional delegations of power to public actors, the Court asks whether the
delegation establishes an "intelligible principle" to which the agency
must conform.[140] One function of the intelligible principle is to enable
judicial review of agency action,[141] meaning the legislative directive must
be "sufficiently definite and precise to enable Congress, the courts, and

seriously confining nondelegation doctrine . . . ."). In his *Whitman* concurrence, Justice
Stevens advocated for a functionalist reframing of this apparent inconsistency, arguing
that "it would be both wiser and more faithful to what we have actually done in delegation
cases to admit that agency rulemaking authority is 'legislative power.' . . . As long as the
delegation provides a sufficiently intelligible principle, there is nothing inherently uncon-
stitutional about it." See 531 U.S. at 488–90 (Stevens, J., concurring in part and concurring
in the judgment).

138. See Yakus v. United States, 321 U.S. 414, 424 (1944) ("The Constitution . . . does
not demand the impossible or the impracticable. . . . The essentials of the legislative
function are the determination of the legislative policy and its formulation and
promulgation as a defined and binding rule of conduct . . . ."); see also *Whitman*, 531 U.S.
at 475 ("[A] certain degree of discretion, and thus of lawmaking, inheres in most execu-
tive or judicial action." (alteration in original) (quoting *Mistretta*, 488 U.S. at 417 (Scalia,
J., dissenting))).

139. See *Whitman*, 531 U.S. at 474 ("[W]e have found the requisite 'intelligible
principle' lacking in only two statutes, one of which provided literally no guidance for the
exercise of discretion, and the other of which conferred authority to regulate the entire
economy on the basis of no more precise a standard than . . . assuring 'fair competition.'"
(citing A.L.A. Schechter Poultry Corp. v. United States, 295 U.S. 495 (1935); Panama
Refining Co. v. Ryan, 293 U.S. 388 (1935))); see also Charles H. Koch, Jr. & Richard
Murphy, 4 Administrative Law and Practice § 11:13 (3d ed. 2019) ("The basic statement of
the law is that a delegation with 'standards' is permissible. However for generations
standards such as 'public interest,' 'reasonable' and 'feasible' have been accepted.").

140. *Whitman*, 531 U.S. at 472 ("[W]hen Congress confers decisionmaking authority
upon agencies *Congress* must 'lay down by legislative act an intelligible principle to which
the person or body authorized to [act] is directed to conform.'" (second alteration in
original) (quoting J.W. Hampton, Jr., & Co. v. United States, 276 U.S. 394, 409 (1928))).
The Court's recent split in *Gundy v. United States* suggests that the intelligible principle test
may not be long for this world. Although Justice Kagan's plurality opinion found that the
delegation at issue "easily passe[d] constitutional muster," 139 S. Ct. 2116, 2121 (2019),
the dissenting Justices criticized the intelligible principle test as "mutated" beyond its
original form, allowing "lawmaking and law enforcement responsibilities [to be] united in
the same hands." Id. at 2138–45 (Gorsuch, J., dissenting, joined by Roberts, C.J. &
Thomas, J.). While Justice Alito provided the decisive vote to uphold the delegation, he
wrote separately to express his willingness to revisit the Court's delegation jurisprudence
in the future. Id. at 2131 (Alito, J., concurring in the judgment) ("If a majority of this
Court were willing to reconsider the approach we have taken for the past 84 years, I would
support that effort. . . . Because I cannot say that the statute lacks a discernable standard
that is adequate under the approach this Court has taken for many years, I vote to
affirm.").

141. Indus. Union Dep't, AFL-CIO v. Am. Petroleum Inst., 448 U.S. 607, 686 (1980)
(Rehnquist, J., concurring in the judgment) (noting that one function of the nondelega-
tion doctrine is to "ensure[] that courts charged with reviewing the exercise of delegated
legislative discretion will be able to test that exercise against ascertainable standards").

the public to ascertain whether the Administrator . . . has conformed to those standards."[142] A delegation's legitimacy therefore depends in part on courts' ability to ensure continued accountability[143]—to understand the legislative intent behind the delegation and determine whether the challenged action conforms to the will of Congress.[144]

While the nondelegation doctrine described above applies the intelligible principle test to delegations of legislative power to public actors, courts have distinguished this formulation from the delegation of power to private actors.[145] Under the private delegation doctrine—the "lesser-known cousin" of nondelegation[146]—"the question . . . becomes whether 'grants of government power to private entities are adequately structured to preserve constitutional accountability.'"[147] This distinction between delegations to public versus private actors is particularly important at a time when government functions are facing increasing privatization.[148]

As courts and scholars have noted, the nondelegation doctrine's concerns about government oversight are heightened when government power is placed in the hands of private actors.[149] Whereas the lax

---

142. *Yakus*, 321 U.S. at 426.

143. See Cass R. Sunstein, Nondelegation Canons, 67 U. Chi. L. Rev. 315, 319 (2000) ("In light of the particular design of the central lawmaking institution, any delegation threatens to eliminate the special kind of accountability embodied in that institution . . . .").

144. *Yakus*, 321 U.S. at 426 ("Only if . . . there is an absence of standards for the guidance of the Administrator's action, so that it would be impossible . . . to ascertain whether the will of Congress has been obeyed, would we be justified in overriding its choice of means for effecting its declared purpose . . . .").

145. See Dep't of Transp. v. Ass'n of Am. R.Rs., 135 S. Ct. 1225, 1237 (2015) (Alito, J., concurring) (explaining that the constitutional justifications for permitting delegations of power within the government do not exist when authority is delegated to private actors). But see Metzger, supra note 120, at 1441 ("[M]any decisions examining private delegations at the federal level use essentially the same framework as is applied to 'public' delegations . . . thereby suggesting that the Court sees such private delegations as presenting nothing beyond ordinary separation of powers issues.").

146. Ass'n of Am. R.Rs. v. U.S. Dep't of Transp. (*Amtrak I*), 721 F.3d 666, 670 (D.C. Cir. 2013), vacated, 135 S. Ct. 1225 (2015).

147. Verkuil, supra note 121, at 89 (quoting Metzger, supra note 120, at 1456).

148. See Kimberly N. Brown, "We the People," Constitutional Accountability, and Outsourcing Government, 88 Ind. L.J. 1347, 1351 (2013) ("Private contractors now perform a broad range of functions for the federal government . . . ."). Compare Verkuil, supra note 121, at 140 (noting that federal contract spending totaled $203 billion in fiscal year 2000), with Contract Explorer, Data Lab (2018), https://datalab.usaspending.gov/contract-explorer.html [https://perma.cc/TP7L-QCZ8] (last visited July 24, 2019) (providing a figure of more than $500 billion for fiscal year 2017).

149. See, e.g., Nat'l Ass'n of Regulatory Util. Comm'rs v. FCC, 737 F.2d 1095, 1143 (D.C. Cir. 1984) (per curiam) ("[T]he difficulties sparked by such allocations [to public actors] are even more prevalent in the context of agency delegations to private individuals."). In this context, "government supervision serves both to assuage concerns regarding the source of lawmaking (and enforcement) authority and to ensure

enforcement of the nondelegation doctrine may be justified on the grounds that the Constitution vests some amount of discretion in the executive branch and provides mechanisms of accountability, a heightened level of scrutiny may be required when these justifications do not exist, as in a private delegation.[150] Although Supreme Court jurisprudence tends to approve of the delegation of government power to private actors, it "emphasize[s] the presence of government review of private decisionmaking in upholding private delegations."[151]

B.   *The History of Private Delegation*

1. *New Deal Origins.* — Although delegations to private actors were typically upheld prior to the mid-twentieth century, the New Deal "gave sharp focus to the private delegation doctrine, as reliance on private regulation and corporatism represented cornerstones of President Roosevelt's early efforts to revive the national economy."[152] In *A.L.A. Schechter Poultry Corp. v. United States,* the Supreme Court considered whether provisions of the National Industrial Recovery Act that permitted private industrial groups to write local trade codes unconstitutionally delegated lawmaking power to those private groups.[153] The majority, led by Chief Justice Hughes, wondered if "it be seriously contended that Congress could delegate its legislative authority to trade or industrial associations or groups so as to empower them to enact the laws they deem to be wise and beneficent."[154] Hughes went on to explain that this delegation could not be "made valid by . . . a preface of generalities as to permissible aims," and that "[s]uch a delegation of legislative power is unknown to our law and is utterly inconsistent with the constitutional prerogatives and duties of Congress."[155]

One year later, the Court decided *Carter v. Carter Coal Co.,* invalidating the Bituminous Coal Conservation Act as an impermissible delegation of power to private parties because it allowed coal miners and producers to enter labor agreements that were binding on all other miners and producers in the local area.[156] Because the Act delegated not

---

transparency and neutrality in the process." Note, The Vagaries of Vagueness: Rethinking the CFAA as a Problem of Private Nondelegation, 127 Harv. L. Rev. 751, 766 (2013) [hereinafter Vagaries].

150. See *Ass'n of Am. R.Rs.*, 135 S. Ct. at 1237 (Alito, J., concurring) ("[T]he Court does not enforce the nondelegation doctrine with more vigilance . . . [because] the other branches of Government have vested powers of their own that can be used in ways that resemble lawmaking. . . . When it comes to private entities, however, there is not even a fig leaf of constitutional justification.").

151. Metzger, supra note 120, at 1439.

152. Id. at 1438 & n.239.

153. 295 U.S. 495, 519–22 & n.4 (1935).

154. Id. at 537.

155. Id.

156. 298 U.S. 238, 311 (1936).

"to an official or an official body, presumptively disinterested, but to private persons whose interests may be and often are adverse to the interests of others in the same business," the Court deemed it "legislative delegation in its most obnoxious form."[157] The Court balked at the idea of Congress conferring power on a private majority "to regulate the affairs of an unwilling minority,"[158] and found that the regulation of coal production "is necessarily a governmental function, since, in the very nature of things, one person may not be entrusted with the power to regulate the business of another."[159]

2. *The Modern Private Delegation Doctrine.* — The Supreme Court has not invalidated a law under the private delegation doctrine since *Carter Coal*,[160] giving the impression that "the most salient characteristic of current private delegation doctrine is its dormant status."[161] Nevertheless, there are indications that the doctrine remains influential, in part because the Supreme Court "has continued to emphasize the presence of government review of private decisionmaking in upholding private delegations."[162]

In assessing a private delegation today, "the pre–New Deal cases remain valid . . . both because they have never been overruled and, more importantly, because the principles on which they relied remain relevant and vital."[163] Despite the dormancy of private delegation at the Supreme Court, both state and lower federal courts have continued to apply the doctrine, helping to provide a fuller picture of its principles and considerations.[164]

---

157. Id.

158. Id.

159. Id.

160. Vagaries, supra note 149, at 764.

161. Metzger, supra note 120, at 1438; see also Michaels, Constitutional Coup, supra note 106, at 125–26 ("[C]ourts have generally declined to treat contractors . . . as the true recipients of delegated powers—and thus subject to the doctrinal bar on private delegations.").

162. Metzger, supra note 120, at 1439. But see Free Enter. Fund v. Pub. Co. Accounting Oversight Bd., 561 U.S. 477, 503 (2009) (explaining that, in the context of removal of government officials, a "standard appropriate for limiting Government control over private [regulatory] bodies may be inappropriate for officers wielding the executive power of the United States").

163. A. Michael Froomkin, Wrong Turn in Cyberspace: Using ICANN to Route Around the APA and the Constitution, 50 Duke L.J. 17, 144 (2000).

164. See Vagaries, supra note 149, at 764 ("[T]he principle against private delegations has by no means been abandoned . . . . Members of the Court have gestured toward the doctrine on multiple occasions, and lower courts deciding cases after *Carter Coal* have fleshed out its requirements."). Likewise, the doctrine "'flourishes . . . in the state courts' applying state constitutional provisions." Id. at 765 n.112 (alteration in original) (quoting Froomkin, supra note 163, at 150); see also David M. Lawrence, Private Exercise of Governmental Power, 61 Ind. L.J. 647, 672 (1986) ("Whatever the federal practice, the state courts continue to actively review private delegations."). Lawrence puts forth several

The recent battle over Amtrak's regulatory capacity illustrates the continued influence of private delegation principles.[165] In 2013, the D.C. Circuit invalidated section 207 of the Passenger Rail Investment and Improvement Act of 2008 (PRIIA), which required the Federal Railroad Administration and Amtrak, a statutorily created, for-profit company, to jointly develop standards and metrics to assess the quality of passenger rail service.[166] In determining that the statute unconstitutionally delegated power to a private actor, Judge Brown emphasized the potential for Amtrak to abuse its statutory authority for private gain,[167] explaining that the doctrine established in *Carter Coal* "ensures that regulations are not dictated by those who 'are not bound by any official duty,' but may instead act 'for selfish reasons or arbitrarily.'"[168] The court also stressed the structural limitations of the private delegation doctrine, rejecting the government's argument that "'[n]o more is constitutionally required' than the government's 'active oversight, participation, and assent' in its private partner's rulemaking decisions."[169]

The Supreme Court later vacated this decision on the grounds that Amtrak is not a private entity,[170] in part because the federal government controls the company's stock and oversees its operations.[171] While Justice Alito agreed with the majority that Amtrak was a government entity, he nevertheless relied on the private delegation doctrine in arguing that the statute may be invalid as a delegation of additional regulatory power to a private arbitrator.[172] Involving a private actor in the regulatory process, Alito argued, would allow the government to "regulate without accountability . . . by passing off a Government operation as an independent private concern."[173] For Congress to delegate to a nongovernmental actor

---

reasons why federal courts take a more restrained approach in applying the private delegation doctrine than state courts, including greater consideration of judicial economy, concerns over federalism, an increased reliance on bright-line rules, and the relative ease of overruling state court decisions through amendment of the state's constitution. See Lawrence, supra, at 672–75.

165. See *Amtrak I*, 721 F.3d 666, 668–69 (D.C. Cir. 2013), vacated, 135 S. Ct. 1225 (2015), remanded to 821 F.3d 19 (D.C. Cir. 2016).

166. See id.

167. See id. at 675 ("[F]undamental to the public-private distinction in the delegation of regulatory authority is the belief that disinterested government agencies ostensibly look to the public good, not private gain.").

168. Id. (quoting Washington ex rel. Seattle Title Tr. Co. v. Roberge, 278 U.S. 116, 122 (1928)).

169. Id. at 673 (quoting Brief for the Appellees at 19, *Amtrak I*, 721 F.3d 666 (No. 12-5204), 2012 WL 5460856).

170. See Dep't of Transp. v. Ass'n of Am. R.R.s., 135 S. Ct. 1225, 1233–34 (2015).

171. See id. at 1232 ("In addition to controlling Amtrak's stock and Board of Directors the political branches exercise substantial, statutorily mandated supervision over Amtrak's priorities and operations.").

172. See id. at 1237 (Alito, J., concurring).

173. Id. at 1234.

would be to violate the carefully constructed system of accountability established by the Constitution.[174]

On remand, while largely ignoring the Supreme Court's holding that Amtrak is a government entity, the D.C. Circuit again invalidated the statute, with only slight modifications to its earlier approach.[175] Rather than holding that the statute improperly delegated to Amtrak as a private entity, the court read *Carter Coal* broadly as prohibiting delegations of authority to *any* self-interested party.[176] The court concluded that, because the PRIIA enabled Amtrak, a self-interested market participant, to regulate its competitors, the statute violated due process.[177] Whereas the Supreme Court majority had focused on the role of government oversight in determining that Amtrak was a government entity,[178] the D.C. Circuit, again relying on *Carter Coal,* found that no level of oversight was sufficient to remedy the conflict of interest, and therefore the due process violation, inherent in Amtrak's regulatory role.[179]

C.   *Principles of Private Delegation*

The D.C. Circuit's obstinacy in light of the Supreme Court's ruling is notably extreme, but the *Amtrak* saga nevertheless demonstrates that private delegation considerations continue to animate courts' assessments of the role of private actors in government. While there is a lack of agreement among courts and scholars as to the exact factors to be considered in evaluating delegations to private actors,[180] or even whether

---

174. Id. at 1237 ("Our Constitution, by careful design, prescribes a process for making law, and within that process there are many accountability checkpoints. It would dash the whole scheme if Congress could give its power away to an entity that is not constrained by those checkpoints." (citation omitted)).

175. Ass'n of Am. R.R.s. v. Dep't of Transp. (*Amtrak II*), 821 F.3d 19, 32–34 (D.C. Cir. 2016) ("We are bound by the Court's conclusion, and we do not disagree with it. . . . But concluding 'Amtrak is not an autonomous private enterprise' is not the same as concluding it is not economically self-interested." (quoting *Ass'n of Am. R.R.s.*, 135 S. Ct. at 1232)).

176. See id. at 27–31 ("Delegating legislative authority to official bodies is inoffensive because we presume those bodies are *disinterested,* that their loyalties lie with the public good, not their private gain.").

177. Id. ("We conclude, as did the Supreme Court in 1936, that the due process of law is violated when a self-interested entity is 'intrusted with the power to regulate the business . . . of a competitor.'" (alteration in original) (quoting Carter v. Carter Coal Co., 298 U.S. 238, 311 (1936))).

178. See *Ass'n of Am. R.Rs.*, 135 S. Ct. at 1231–32.

179. *Amtrak II*, 821 F.3d at 34 (explaining that, in *Carter Coal,* "what was offensive about the statute was its 'attempt[] to confer' the 'power to regulate the business of . . . a competitor,'" and that "government oversight would not have cured a grant of regulatory power antithetical to the very nature of governmental function" (first alteration in original) (quoting *Carter Coal,* 298 U.S. at 311)).

180. Compare Vagaries, supra note 149, at 764–65 (citing three factors—"first, 'whether [the delegation] authorizes private actors to make law in a non-neutral,

delegation to private actors is permissible under any circumstances,[181] common themes emerge from the jurisprudence: emphasis on govern-ment oversight and a desire to limit potential conflicts of interest. At the heart of both concerns is the need to maintain the role of the Constitution, and therefore the public, in directing the exercise of government power.[182]

1. *Limiting Conflicts of Interest.* — A central concern in private delegations is that "private entities may face significant conflicts of interest or other tensions with public goals as a result of market incentives or professional culture."[183] This idea played a key role in *Carter Coal*, in which the Court objected to the ability of private parties in the coal industry to "regulate the business of another, and especially of a competitor."[184] The potential for conflicts of interest in government decisionmaking has animated Supreme Court jurisprudence in other areas of the law as well.[185] For example, in *Gibson v. Berryhill*, the Court prohibited the Alabama Board of Optometry from adjudicating delicens-ing proceedings within the optometry industry because individual board members, who were also private practitioners, had a financial interest in

---

nontransparent way'; second, 'whether affected parties are adequately represented in the private lawmaking process'; and third, 'whether the state retains control over the private delegate'" (quoting David Horton, Arbitration as Delegation, 86 N.Y.U. L. Rev. 437, 474 (2011))), with Froomkin, supra note 163, at 28 ("[T]he private nondelegation doctrine focuses on the dangers of arbitrariness, lack of due process, and self-dealing when private parties are given the use of public power.").

181. Compare *Amtrak I*, 721 F.3d 666, 670 (D.C. Cir. 2013) ("Federal lawmakers cannot delegate regulatory authority to a private entity. To do so would be 'legislative delegation in its most obnoxious form.'" (quoting *Carter Coal*, 298 U.S. at 311)), with Pittston Co. v. United States, 368 F.3d 385, 394 (4th Cir. 2004) ("Not every Congressional delegation of authority to a private party is impermissible . . . .").

182. See Verkuil, supra note 121, at 15–16 ("The People is the sovereign and the Congress and president are her agents. . . . Thus when the Congress . . . delegates to private parties, the Constitution still umpires the relationships. . . . When sovereign powers are delegated, it is with the permission of the People.").

183. Nina A. Mendelson, Supervising Outsourcing: The Need for Better Design of Blended Governance, *in* Administrative Law from the Inside Out 427, 432–33 (Nicholas R. Parrillo ed., 2017) [hereinafter Mendelson, Supervising Outsourcing].

184. *Carter Coal*, 298 U.S. at 311.

185. See Gibson v. Berryhill, 411 U.S. 564, 578 (1973); see also Connally v. Georgia, 429 U.S. 245, 250–51 (1977) (per curiam) (finding a due process violation in the issuance of a search warrant by a justice of the peace who received a fee when warrants were issued but not when they were denied); Tumey v. Ohio, 273 U.S. 510, 523 (1927) ("[I]t certainly violates the Fourteenth Amendment, and deprives a defendant in a criminal case of due process . . . [when] the judge . . . has a direct, personal, substantial, pecuniary interest in reaching a conclusion against him in his case."); Verkuil, supra note 121, at 142 ("Applied due process challenges based on biased private adjudicators . . . continue to be successful in federal and state courts.").

the adjudications.[186] Although the Board itself was a public body, individual members stood to gain by delicensing optometrists who might compete with their own private businesses,[187] leading the Court to conclude that "those with substantial pecuniary interest in legal proceedings should not adjudicate these disputes."[188]

The concern that conflicts of interest will prevent private actors from wielding government power in a neutral way was echoed nearly eighty years after *Carter Coal* in the *Amtrak* case, when the D.C. Circuit expressed concern that the power given to Amtrak allowed self-interested motives to overshadow consideration of the public good in the regulatory process.[189] Even on remand, after the Supreme Court held that Amtrak was not a private entity for the purposes of nondelegation, the D.C. Circuit stressed the self-interested nature of Amtrak's regulatory power in invalidating the statute.[190] As noted above, the court's intransigence was extreme, but these opinions nevertheless emphasize the potential for "[s]kewed incentives" that may lead a private delegate to abuse public authority for personal gain.[191]

2. *Ensuring Government Oversight.* — A second theme that emerges from both the early Supreme Court cases and more recent lower court cases is that delegations of authority to private actors may be constitutional when mechanisms for government oversight of private decisions are in place.[192] As Verkuil explains, "When powers are delegated . . . Congress has a constitutional stake in the process. Its job is to assure the

---

186. See 411 U.S. at 578 ("[T]he Board's efforts would possibly redound to the personal benefit of members of the Board, sufficiently so that . . . the Board was constitutionally disqualified from hearing the charges . . . .").

187. See id. at 567–71.

188. See id. at 579.

189. *Amtrak I*, 721 F.3d 666, 675 (D.C. Cir. 2013), vacated, 135 S. Ct. 1225 (2015) ("[F]undamental to the public-private distinction in the delegation of regulatory authority is the belief that disinterested government agencies ostensibly look to the public good, not private gain. For this reason, delegations to private entities are particularly perilous.").

190. See supra notes 175–179 and accompanying text.

191. *Amtrak I*, 721 F.3d at 677.

192. See supra note 151 and accompanying text; see also Vagaries, supra note 149, at 769 ("[L]ower courts upholding delegations have emphasized how private activities can be rendered merely advisory or ministerial when exercised under 'pervasive surveillance and authority' of government officials." (quoting United States v. Frame, 885 F.2d 1119, 1129 (3d Cir. 1989))). But see supra note 179 and accompanying text (explaining the D.C. Circuit's insistence that no level of government oversight is sufficient to cure the conflict of interest inherent in Amtrak's authority to regulate its competitors). For the argument that judicial oversight of agency action is essential to the separation of powers analysis at the heart of the nondelegation doctrine, see Cynthia R. Farina, Statutory Interpretation and the Balance of Power in the Administrative State, 89 Colum. L. Rev. 452, 487–88 (1989) ("A crucial aspect of the capacity for external control upon which the permissibility of delegating regulatory power hinged was judicial policing . . . . The constitutional accommodation ultimately reached in the nondelegation cases implied that principal power to say what the statute means must rest *outside* the agency, in the courts.").

People that the work of government stays in the hands of those responsible for its execution."[193]

One function of government oversight under the private delegation doctrine is to ensure that private parties act within the scope of their delegated power.[194] In this way, the role of oversight in a private delegation aligns with the intelligible principle standard, which functions in part to enable judicial review of an agency's compliance with the legislative will.[195] By requiring the delegating statute to provide an oversight mechanism, the private delegation doctrine emphasizes the government's role in enforcing the boundary between the proper use of authority and a deviation from the delegated power. A secondary function of government oversight may be to monitor private delegations for the conflicts of interests discussed above[196]—a principle illustrated by *Geo-Tech Reclamation Industries, Inc. v. Hamrick*.[197] In *Hamrick*, the Fourth Circuit explained that it was "unable . . . to discern within the language [of the statute] any meaningful standard" to oversee the actions of the private party, allowing administrative decisionmaking to be "subservient to selfish or arbitrary motivations."[198] By this reasoning, government supervision of private delegations accomplishes two things: It allows a public actor to review the delegate's actions for compliance with the legislature's intent and, through this review, limits the private actor's ability to use its public power for private ends.

The emphasis on government supervision may also serve to maintain a formal government presence in the private exercise of government power. While a private actor may be tasked with administrative and advisory duties,[199] key decisions about the policy and structure of the

---

193. Verkuil, supra note 121, at 194; see also Metzger, supra note 120, at 1457 (discussing the "insight that the structure of a private delegation should matter more in determining its constitutionality than the mere fact that private actors are exercising government power").

194. See, e.g., Thomas v. Union Carbide Agric. Prods. Co., 473 U.S. 568, 592–93 (1985) (noting that a statutory provision permitting judicial review of the decision of a private arbitrator "protects against arbitrators who abuse or exceed their powers or willfully misconstrue their mandate under the governing law"); Todd & Co. v. SEC, 557 F.2d 1008, 1014 (3d Cir. 1977) ("The independent review function entrusted to the S.E.C. is a significant factor in meeting serious constitutional challenges to this self-regulatory mechanism.").

195. See supra notes 140–144 and accompanying text.

196. See supra section III.C.1.

197. 886 F.2d 662 (4th Cir. 1989). In *Hamrick*, the Fourth Circuit considered whether West Virginia's Solid Waste Management Act, under which the operation of a landfill required a permit that could be denied based on negative public comments, was an unconstitutional delegation of power to "a narrow segment of West Virginia's citizenry," who could effectively block an application from being approved. See id. at 663–64.

198. Id. at 666–67.

199. See supra note 192 and accompanying text.

delegation must be decided and explained by the legislature.[200] This idea is reflected in the Supreme Court's New Deal discussions of private delegation as well. Four years after *Carter Coal*, the Court upheld a later incarnation of the Bituminous Coal Act, which gave the National Bituminous Coal Commission the power to set minimum coal prices after considering proposals made by local boards of coal producers.[201] The statute provided that these proposals could be "approved, disapproved, or modified by the Commission," and allowed the local boards to serve "as an aid to the Commission but subject to its pervasive surveillance and authority."[202] In short, the Court determined that because the local boards "function[ed] subordinately to the Commission" and were not entrusted with the power to set prices, there was no improper delegation.[203] Private actors may therefore play a role in government, but oversight mechanisms cannot be superficial, and a government actor must be actively involved in evaluating and moderating the decisions of the private delegate.[204]

IV. CREATING CONSTITUTIONAL ACCOUNTABILITY IN RISK ASSESSMENT

The use of privately developed risk assessment tools in sentencing presents the same concerns that the private delegation doctrine is intended to address: a lack of government oversight and the potential for private self-interest to overshadow the public good.[205] While a court is unlikely to strike down the use of these tools in sentencing as an impermissible delegation,[206] private delegation principles can provide a useful

---

200. See Verkuil, supra note 121, at 142 (explaining that in *Carter Coal* "private parties in effect implemented government policy without the approval of public officials; those sworn to uphold the Constitution were excluded from the decision process"); see also Gen. Elec. Co. v. N.Y. State Dep't of Labor, 936 F.2d 1448, 1455 (2d Cir. 1991) ("[A] legislative body may not constitutionally delegate to private parties . . . without supplying standards to guide the private parties' discretion."). In *General Electric*, the Second Circuit struck down a New York labor law that gave the Department of Labor discretion in setting wage rates based on privately negotiated collective bargaining agreements because, in practice, the Department did not exercise this discretion. Its reliance on the privately negotiated rates "constitute[d] . . . an unconstitutional delegation of legislative authority." See id. at 1458.

201. Sunshine Anthracite Coal Co. v. Adkins, 310 U.S. 381, 387–88 (1940).

202. Id. at 388.

203. Id. at 399.

204. See Mendelson, Supervising Outsourcing, supra note 183, at 429 ("[W]e cannot readily assume that an agency official's simple presence will suffice to avoid privatization's dysfunctions and to supply public accountability. More careful attention to how agencies supervise outsourcing—and how they *should* conduct such oversight—is clearly needed.").

205. See supra section II.A (discussing the lack of transparency surrounding the decisions made by private actors in developing risk assessment algorithms, as well as the difficulty that judges may encounter in attempting to understand and apply the results of these algorithms at sentencing).

206. See supra note 161 and accompanying text.

framework for understanding the constitutional accountability gap that arises from the use of these tools.[207] Framing the work of risk assessment developers in private delegation terms, this Part proposes legislative remedies to close the accountability gap created by the use of algorithmic tools in sentencing. Section IV.A identifies oversight failures and conflicts of interest in the use of privately developed risk assessment instruments as the source of the accountability gap discussed in section II.C. Section IV.B then proposes legislative remedies to mitigate these issues, restoring accountability to the use of these technologies.

A.   *Applying Private Delegation Principles to Risk Assessment*

As discussed in Part II, judges lack the necessary information to review and apply risk assessment algorithms in a critical and meaningful way, giving the private developers of these tools a significant role in sentencing determinations.[208] This section applies a private delegation lens to this problem, concluding that the sparse statutory regimes that govern risk assessment provide limited oversight and fail to check conflicts of interest, giving rise to the accountability gap discussed in section II.C.

1. *Potential for Conflicts of Interest.* — On the surface, the potential for private interests to overshadow the public good in the risk assessment context appears minimal because the incentives of private risk assessment developers align with those of the government actors applying their tools. Judges want to accurately predict an offender's recidivism risk to ensure the efficient use of government resources,[209] minimize the risk of future crime,[210] and protect their own political reputations.[211] Likewise, risk assessment developers are motivated to produce accurate tools to increase the demand for their products. In this sense, the public and private incentives are closely aligned, perhaps assuaging concerns that risk assessment developers may abuse the power that is delegated to them.

But the private motives of algorithm developers nevertheless influence the way that these tools are used. As noted in section II.A, private companies may rely on trade secret protections to prevent disclosure of information about their software out of concern that competitors will exploit this information for competitive advantage.[212] Furthermore, some have suggested that developers avoid disclosure not solely out of concern about misappropriation of their intellectual property but also to avoid added scrutiny of their product that might negatively impact their busi-

---

207. Metzger, supra note 120, at 1394–410.
208. See supra section II.B.
209. See supra note 19 and accompanying text.
210. See supra note 21 and accompanying text.
211. See supra note 95 and accompanying text.
212. See supra notes 74–79 and accompanying text.

ness.[213] Likewise, during design, developers may opt for a machine learning model that is more likely to produce false positives than false negatives,[214] since recidivism by a "low risk" offender after release is worse for business than the imprisonment of someone deemed "high risk" who actually would not have reoffended.[215] Private motivations therefore have the potential to influence how risk assessment algorithms function, as well as how judges are able to interpret their results, to the detriment of the public good.

2. *The Lack of Government Oversight.* — The private delegation lens also reveals clear gaps in government oversight of the companies developing risk assessment tools. Often, the statutory frameworks requiring or permitting the use of risk assessment tools in sentencing are sparse and provide limited instruction on how risk scores should be applied.[216] Because of legal and technological opacity,[217] judges are unable to review the choices made by private developers to ensure that a given risk assessment algorithm functions in a way that is consistent with the purpose and scope of the delegation,[218] or even with established sentencing policy.[219] Furthermore, existing statutory frameworks often provide no instruction to tool developers on how to construct risk assessment tools, leaving these private actors to make their own policy determinations when gathering data, defining recidivism, and developing an algorithm.[220] At the same time, these statutes often provide little direction to judges on how to apply the risk score they receive, but judges have such limited information about the meaning of the risk score in front of them that it may be difficult to exercise this discretion in a meaningful way.[221]

---

213. See supra note 97 and accompanying text. For an overview of the way that the criminal justice system incentivizes both public and private institutions to perpetuate overcriminalization, see generally Eisha Jain, Capitalizing on Criminal Justice, 67 Duke L.J. 1381 (2018).

214. See supra note 48 and accompanying text.

215. Cf. Kehl et al., supra note 15, at 14 (explaining that a judge "may simply take a risk-averse approach and impose more stringent sentences on criminals who are labeled high risk in order to avoid potential blame for a high-risk criminal who received a less severe sentence and ultimately did reoffend").

216. See supra section I.A (providing several examples of statutes permitting or requiring the use of risk assessment tools in sentencing); see also infra note 223.

217. See supra section II.A.1.

218. See supra section II.B.

219. See supra note 116 and accompanying text (explaining that tool developers often fail to consider existing legal limitations on the use of certain predictive factors in sentencing).

220. See supra notes 108–117 and accompanying text (identifying points throughout the process of creating a risk assessment algorithm when developers make decisions that can influence a tool's outcomes).

221. Cf. Mendelson, Private Control, supra note 65, at 784–85 (explaining that, in the context of incorporation of private standards by reference, agency oversight is inhibited by

The existing statutory frameworks permitting or requiring the use of predictive risk assessment tools create an overly broad and undefined delegation of power to private actors in sentencing.[222] To give judges a meaningful opportunity to review risk scores, tool vendors must provide additional information about how their products are developed and function. Likewise, more information is needed from legislators about the purpose and policies of using risk assessments in sentencing, enabling courts to review the actions of private developers for conformity and to apply risk assessment results in a way that promotes meaningful judicial discretion.[223]

---

increasing reliance on private expertise, which can reduce the cost of developing standards but may also decrease agencies' ability "to fully evaluate whether the privately developed standard meets public goals").

222. While at least one circuit has noted that "[e]ven an intelligible principle cannot rescue a statute empowering private parties to wield regulatory authority," *Amtrak I*, 721 F.3d 666, 671 (D.C. Cir. 2013), vacated, 135 S. Ct. 1225 (2015), the search for an intelligible principle in the use of risk assessment algorithms in sentencing sheds light on just how little control the legislature exercises in these instances. In *Mistretta v. United States*, the Court considered whether the Sentencing Reform Act (SRA), which created the United States Sentencing Commission and enabled the Commission to create sentencing guidelines for all federal crimes, was an unconstitutional delegation of legislative power. 488 U.S. 361, 367–71 (1989). Ultimately, the Supreme Court upheld the SRA's grant of authority to the Sentencing Commission because the Act provided an intelligible principle to limit the Commission's actions by giving policy objectives and detailed explanations of the Commission's duties and methodology for developing sentencing guidelines. Id. at 370–79.

A comparison of the SRA and the state statutes providing for the use of risk assessment algorithms in sentencing reveals a deep disparity in the level of guidance provided by the legislatures. Whereas the SRA includes a clear statutory purpose, numerous policy goals, and enumerated duties for the Sentencing Commission, all of which contributed to the Court's finding of an intelligible principle in *Mistretta*, the state statutes discussed in section I.A often provide little guidance beyond the requirement or suggestion that risk assessment tools be incorporated into a judge's sentencing decision. See supra notes 27–32 and accompanying text. For a judge presented with an offender's risk score at sentencing, a state procurement officer seeking to license a risk assessment tool, or a private company beginning to develop a predictive algorithm, these statutes offer limited information on the legislatures' goals and policies to guide the creation and use of these instruments.

223. See supra note 98 (discussing the ambiguity surrounding how risk assessment algorithms are intended to be used and the difficulty of determining how often a judge should deviate from an algorithm's recidivism risk determinations); see also Verkuil, supra note 121, at 141 ("Often the contractual delegations outsource services that transfer authority to private hands without adequate descriptions of the services to be provided. When that happens, the danger that government functions involving 'significant authority' are also transferred is heightened and control of the public enterprise is threatened."); John D. Donahue, The Transformation of Government Work, *in* Government by Contract: Outsourcing and American Democracy 41, 45 (Jody Freeman & Martha Minow eds., 2009) ("To outsource a function you not only need to be able to say what you want[,] . . . you also need to be in a position to know what you've gotten . . . . The easier it is to monitor performance . . . , the more safely a task can be delegated.").

B. *Restoring Accountability and Oversight*

Building on section IV.A's identification of oversight failures and conflicts of interest in the use of privately developed risk assessment algorithms, this section proposes legislative solutions to close the accountability gap arising from the use of these tools. Legislative and regulatory changes, as opposed to litigation, are likely to be effective means of increasing accountability in the use of privately developed risk assessment algorithms. Although enforcement of the private delegation doctrine may be more common at the state level,[224] courts are still reluctant to invalidate existing delegations to private entities.[225] In the context of privately developed risk assessments, in which the complexity and private nature of the technology limit government oversight and control, the enforcement of accountability through judicial review is reduced.[226] While a private delegation challenge is unlikely to succeed in court, the doctrine highlights the legislature's policymaking role and provides important principles for developing legislative remedies. State legislatures and sentencing commissions often have the resources to develop comprehensive, informed solutions,[227] whereas courts may not be able to craft widely applicable, nuanced reforms within the confines of a single case.[228]

To restore constitutional accountability to the algorithm-assisted sentencing process, the structure of the laws permitting the use of this technology must be modified.[229] As Verkuil explains, when "making

224. See supra note 164 and accompanying text.

225. See Michaels, Constitutional Coup, supra note 106, at 126; see also Verkuil, supra note 121, at 143 (describing the risks of Lochnerism in broad judicial enforcement of the bar on delegations to private actors).

226. See Mendelson, Supervising Outsourcing, supra note 183, at 442–43 (explaining that external accountability mechanisms such as judicial review "may be less useful to address inadequate agency supervision of outsourced activity"); cf. Mendelson, Private Control, supra note 65, at 789 (explaining that lack of access to information about privately developed rules inhibits the effectiveness of accountability mechanisms).

227. See infra note 234 and accompanying text.

228. For example, in *State v. Loomis*, the Wisconsin Supreme Court listed several warnings that should be provided to judges applying the COMPAS tool in the future, including that Northpointe does not disclose what it considers to be proprietary information about the tool and that studies suggest that the tool disproportionately categorizes minorities as high risk. 881 N.W.2d 749, 769 (Wis. 2016). Many have questioned the efficacy of these warnings, in part because they do not address the root problems that arise from the use of risk assessment tools. See, e.g., Recent Case, supra note 22, at 1531 ("The court's advisement . . . ignores judges' inability to evaluate risk assessment tools, and it fails to consider the internal and external pressures on judges to use such assessments."); Jones, supra note 19 ("[C]ritics contend that such warnings, when weighed against the gloss of objectivity provided by data, will amount to little true scrutiny of these tools.").

229. Other possible ways to increase available information about risk assessment tools include modifications to state freedom of information laws, see Carlson, supra note 14, and limitations on trade secret privileges in the criminal context, see Wexler, supra note 70, though none of these modifications directly addresses the constitutional accountability

public actions 'private,' . . . [delegations] should come with strings attached that ensure fairness at the individual level and accountability at the political level."[230] Because of its focus on the structure of the relationship between public and private actors, the private delegation doctrine provides a useful framework for designing legislative and regulatory remedies that address the need for increased information from both risk assessment developers and legislators.[231] Even without expertise in statistics or computer science, thoughtful legislators can reshape the way in which risk assessment algorithms are used and restore accountability through more purposive and deliberate lawmaking.[232]

To increase constitutional accountability in this delegation, state statutes should require sentencing commissions to develop additional guidelines and reporting requirements for the use of risk assessment algorithms in sentencing.[233] Sentencing commissions should identify the dataset from which a tool will be developed, determine the definition of recidivism to be used in training the algorithm, and set a minimum acceptable error rate for recidivism risk prediction. Sentencing commissions may be particularly well-positioned to carry out these policymaking

---

problem discussed in Part II. While these solutions may increase access to a tool's source code, they would likely require case-by-case litigation and may not produce information that an average offender or judge would be able to interpret.

230. Verkuil, supra note 121, at 80–81; see also Metzger, supra note 120, at 1394–95 ("[P]rivatization is poorly characterized as government withdrawal or disinvolvement from an area of activity. . . . Rather than government withdrawal, the result is a system of public-private collaboration . . . in which both public and private actors share responsibilities.").

231. See supra note 193 and accompanying text (noting the Supreme Court's emphasis on the structure of private delegations in assessing their constitutionality).

232. See David B. Spence, Managing Delegation Ex Ante: Using Law to Steer Administrative Agencies, 28 J. Legal Stud. 413, 415 (1999) ("[W]hile Congress cannot foresee many of the important policy decisions it delegates to the agency, it can use enabling legislation to shape the agency policy-making process in ways that influence subsequent agency policy decisions."). However, the failure of New York's algorithmic accountability law in 2017 demonstrates the shortcomings of the political branches in insisting on constitutional accountability from private companies. See Julia Powles, New York City's Bold, Flawed Attempt to Make Algorithms Accountable, New Yorker (Dec. 20, 2017), https://www.newyorker.com/tech/annals-of-technology/new-york-citys-bold-flawed-attempt-to-make-algorithms-accountable [https://perma.cc/DFJ8-9HKA] (explaining that the bill was significantly watered down after being met with "strong resistance from . . . tech companies, which argued that it would force them to disclose proprietary information, supposedly harming their competitive advantage").

233. There may be some variation among states in the appropriate government body to develop these guidelines and requirements, as the mandate and authority of sentencing commissions can vary substantially from state to state. See Neal B. Kauder & Brian J. Ostrom, Nat'l Ctr. for State Courts, State Sentencing Guidelines: Profiles and Continuum 7–27 (2008), https://www.ncsc.org/~/media/Microsites/Files/CSI/State_Sentencing_Guidelines.ashx [https://perma.cc/KEL9-GSDG] (providing an overview of the purpose, membership, and work of several state sentencing commissions). At the very least, the power to set key policies for the development and use of risk assessment tools should lie with a government actor, whether it be the state legislature or sentencing commission.

tasks, as their membership typically includes judges, district attorneys, public defenders, academic experts in criminal justice, and community members.[234] By requiring a public actor to make these key decisions, legislators can ensure that critical policy decisions remain in the hands of publicly accountable officials and provide judges with greater transparency around the choices that shaped the algorithm.

In addition to prescribing policies and methodologies for the development of these algorithms, risk assessment statutes should also direct sentencing commissions to establish reporting procedures for tool developers, requiring them to provide error rates, discrepancies in accuracy among different populations, comparative rates of false positives and negatives, and other basic information about the performance of their algorithms.[235] By requiring this reporting, sentencing commissions can provide judges with necessary context for understanding both an offender's risk score and the broader limitations of algorithmic prediction. With this information readily available, sentencing judges may be less inclined to rely heavily on algorithmic outputs, developers may be faced with fewer demands for proprietary source code, and other government actors may be better equipped to assess the efficacy of these tools. In returning key policymaking functions to a government body and enabling judges to better understand the function and limitations of risk assessment tools, legislators can ensure that adequate standards are in place to guide tool developers[236] and that public actors maintain "surveillance and authority" over the actions of these private parties.[237]

CONCLUSION

As the applications of machine learning expand and government willingness to contract with private developers of algorithmic risk assessment tools increases, it is important to take stock of the accountability

---

234. See id. (identifying the required members of several state sentencing commissions).

235. The Idaho legislature recently amended the state's criminal procedure law to require that "[a]ll documents, data, records, and information used by the builder to build or validate the pretrial risk assessment tool and ongoing documents, data, records, and written policies outlining the usage and validation of the pretrial risk assessment tool" be made publicly available, and to grant defendants in criminal proceedings the right to "review all calculations and data used to calculate the defendant's own risk score." See Idaho Code § 19-1910 (2019). While this is certainly a step toward meaningful disclosure, the statute notably fails to provide a role for public officials in making the policy decisions that form the basis of the risk assessment tool, instead leaving those tasks to an unidentified "builder."

236. See supra notes 194–198 and accompanying text (explaining the importance of meaningful standards to guide government review of private actions).

237. See supra notes 199–204 and accompanying text (highlighting the maintenance of government supervision and control of private actors as a significant concern in private delegation cases).

issues that arise from the public use of private technology. To do so, this Note explores the role that private actors, who operate outside traditional mechanisms of constitutional accountability, play in shaping the outcomes these tools produce. The private nature of many recidivism risk assessment algorithms leaves sentencing judges unable to understand and adequately apply their results, leading to a greater reliance on the policy decisions of private developers. As a result, private actors are given an outsized role in sentencing, and the legislative purpose of risk assessment statutes is undermined. To remedy this problem, this Note uses the private delegation doctrine as a framework to improve judicial engagement with algorithmic risk scores. Because existing statutory frameworks allow private actors to wield government power with limited public oversight and control, legislators must increase the specifications included in risk assessment statutes, bolstering the ability of judges to understand and apply these technologies.