MAY 2020

KRISTIN PORTER
(MDRC CENTER FOR DATA INSIGHTS)

CINDY REDCROSS
(MDRC CENTER FOR CRIMINAL JUSTICE RESEARCH)

LUKE MIRATRIX
(HARVARD UNIVERSITY)

BALANCING PROMISE AND CAUTION IN PRETRIAL RISK ASSESSMENTS

INTRODUCTION

Across the country, release and detention decisions for defendants in the pretrial period — that is, the period after arrest while a criminal case is being adjudicated — are increasingly guided by risk assessments, which rely on data to estimate defendants' risk of failing to appear for a court date or of being charged with new criminal activity if released pending trial. A risk assessment is generally used by a judicial body to help determine whether a defendant will be released while waiting for a case to be resolved, and if so, under what conditions (these are the defendant's *release conditions*). The goal of pretrial risk assessments is to make any restrictions imposed on a defendant's liberty better align with the risk that person poses to the community.

As has been widely discussed, these assessments — and the way they are used — have the potential to introduce new biases and further perpetuate racial disparities in the criminal justice system.[1] This research brief describes the approach taken by MDRC's Center for Criminal Justice Research and MDRC's Center for Data Insights to understand, assess, and address bias in pretrial risk assessments and the systems that use them.

First and foremost, to understand the ways pretrial risk assessments can affect defendants' release outcomes, one must understand the context in which those assessments are used. Risk assessments are typically designed to provide *additional* information to judges and others making release decisions. They are typically used in conjunction with jurisdiction-specific decision-making frameworks that use a defendant's risk score in combination with local statutes and policies to produce a recommendation for release conditions. In the end, these recommendations are suggestions, and the decision-making body takes them into account along with other factors.

Risk assessments also tend to be adopted for the first time as part of larger pretrial justice system reform efforts. And they are generally adopted as guidelines, with enormous latitude in how they are used. As a result, while knowing how a specific tool is biased is an important part of the puzzle, it does not give the entire picture. Researchers need to evaluate whole systems, not just single risk-assessment tools in isolation. They should examine, for example, whether there are disparate impacts on actual release outcomes for different races and other groups. Researchers could also assess impacts on existing disparities, for

mdrc
BUILDING KNOWLEDGE TO IMPROVE SOCIAL POLICY

1 Among many journal articles and other discussions, see Angwin, Larson, Mattu, and Kirchner (2016); Calabresi (2014); Klingele (2015); Stevenson (2019); Kleinberg et al. (2017); Corbett-Davies et al. (2017).

example by measuring the effects of an overall policy initiative on the ratios or gaps in detention between racial groups.

As an example, an MDRC study in Mecklenberg County, North Carolina (the county that contains Charlotte) found that the adoption of pretrial risk assessments increased the percentage of defendants who were released pending trial. It also found no evidence that the risk-assessment policy changes affected the percentages of defendants who made all of their court appearances or who were charged with new crimes while waiting for their cases to be resolved — even though a higher proportion of defendants in the system faced felony charges in the period after the risk assessment was implemented.[2] In this context, most of the changes in pretrial release conditions occurred at a step in the pretrial case process before the risk-assessment report was completed. Thus, the information in the risk assessment itself could have had at most only a small effect on the way judges set release conditions. It appears the overall reform, not the specific tool, was responsible for most of the changes.

Meanwhile, MDRC found no evidence of racial disparity in the way judges set release conditions when accounting for risk level; specifically, bail setting and detention did not differ between Black and White defendants within levels of low, medium, and high risk. There was also no evidence that the risk assessment had any effect on racial disparities in detention — meaning that there was no improvement or worsening of disparate rates of detention.

Differences in group behavior and treatment make it particularly difficult to assess bias and impacts on bias. For example, in Mecklenburg County, Black defendants were more likely than other racial groups to be assessed as being at high risk. It is impossible to know to what degree Black defendants are truly at higher risk in Mecklenburg County or to what degree historical biases in the criminal justice data used in the algorithm led to higher risk scores for those Black defendants. Regardless, this discrepancy in assessed risk levels for racial groups meant that it would be misleading to evaluate the impacts of the risk-assessment policies on racial disparities without accounting for differing risk levels.

Similarly, in an evaluation of New Jersey's sweeping bail reform, MDRC found that impacts were smaller in magnitude for Black defendants.[3] However, in order to fully understand how New Jersey's bail reform affected racial disparity in outcomes, one must here account for differences in the types of charges incurred among the racial groups, especially given that the impacts of the bail reform differed substantially by charge class.

In reflecting on these examples, several mechanisms for bias become evident: There can be bias in the tool (the usual concern raised when reflecting on risk-assessment tools), bias in the way the tool is implemented (even with an unbiased tool, the way the tool is implemented can introduce bias — or

2 Redcross and Henderson (2019).

3 Golub, Redcross, and Valentine (2019).

could offset bias in the tool), and finally bias in the *evaluation* of the system or tool. It is worth highlighting this last aspect because a preexisting bias could be propagated by a risk-assessment tool and then not be detected in an evaluation of that tool. All these mechanisms for bias can damage the tool along with researchers' understanding of what it is doing.

Because there are multiple mechanisms for bias, one must acknowledge that those tools are generally parts of broader reforms in order to understand their impact. The tools need to be evaluated as they are being used. At minimum, researchers should assess both the risk-assessment scores themselves and the recommendations from local decision-making frameworks that use the scores. Ideally, studies should also examine the actual decisions made as those frameworks are applied in the real world.

Traditional fairness work has focused primarily on the risk assessment itself, and this piece is obviously incredibly important. But much of this work has also operated in a context where some forms of bias are not typically accounted for — possible bias in the measured outcome, for example. That being said, the algorithmic fairness literature has also highlighted serious concerns, such as how different stakeholders can be interested in different measures of fairness that are fundamentally not reconcilable. Policy evaluations must take these lessons on bias one step further.

Another concern illuminated by the examples above is the importance of viewing reform as an attempt to improve a system that is in place. Given the multiple mechanisms for potential bias, one should perhaps focus on change and improvement in a policy reform. An evaluation should compare the bias in the new system with the bias in the old. A good evaluation of a successful reform could both establish some degree of improvement and also highlight opportunities for future change.

This paper is organized as follows: It first identifies and explains different ways bias could arise in a system that includes a risk assessment. It then discusses different strategies for assessing bias, and finally touches on methods for addressing it. Overall, it advocates thinking systematically about the many sources of bias and how they can affect a risk-assessment tool, its implementation, and its overall impact. The paper's goal is to shed light on methodological approaches for assessing and addressing bias both in risk-assessment tools and in how they are used, the latter of which is often not often addressed in research on the tools. It argues that while it is important to assess and address bias in pretrial risk-assessment tools, it is also essential to assess and address bias in their implementation. The primary goal should be to evaluate the overall and disparate impacts of the reform efforts that are using these tools.

SOURCES OF BIAS

In general, pretrial risk assessment involves three steps: (1) use a defendant's demographic and criminal history to assess risk, (2) interpret the risk in order to recommend the conditions under which that defendant should be released, and then (3) execute a final decision based on the risk estimate, the jurisdiction's statutory requirements, the recommended decision, other evidence (for example, facts about the case), and other factors (for example, judges' decision-making). In each step, there is the potential for bias. This section identifies several sources of potential bias and where bias can affect this process. The following sections then present ideas for how researchers can assess and address such bias.

For simplicity, this paper focuses on bias that can affect Black defendants, although any of the ideas discussed here could be applied to other demographic subgroups.

1 **Bias in the data.** In the administrative data used to develop and validate pretrial assessments, some data elements are very likely to be inherently biased. Consider measures of criminal activity, for example. One factor leading to recorded criminal activity is the likelihood of getting arrested. This likelihood can vary by race even when the true level of criminal activity is constant. For example, individuals who live in areas with elevated amounts of policing (for example, "stop and frisk") will tend to have more arrests than individuals with the same criminal behavior who live in areas with less policing. The level of policing in neighborhoods is often correlated with race. Therefore, White and Black populations with the same actual prevalence of criminal behavior will be arrested at different rates for such behavior.[4]

The resulting bias in the data affects both the risk factors used as predictors in a risk assessment and the outcomes predicted by a risk assessment. An assessment that depends on risk factors that are biased (for example, past criminal activity) may perpetuate this bias because it could score a Black defendant as at higher risk than a White defendant with the same true level of risk. With respect to outcomes, particularly the outcome of new criminal activity, Black defendants awaiting trial and living in highly policed areas may be more likely to be picked up for similar offenses than White defendants awaiting trial in other areas. Therefore, even if the risk factors in a risk assessment are not biased, the evaluation of the risk assessment could suggest bias, as a larger proportion of Black defendants with the same risk score could have new arrests than did similar White defendants. Additionally, if a risk assessment were calibrated to such data, to obtain similar rearrest rates for subgroups within a given estimated risk, then the resulting model could systematically assign higher risk scores to what are truly lower-risk Black defendants.

2 **Bias in a given predictive model.** Even if the risk factors and outcomes involved in a risk assessment are all measured without bias, the underlying model can still be biased. This situation can occur if the relationships between the risk factors and the outcome are different for different racial groups. Take, for example, a condition where prior arrests predict an outcome better for White defendants than Black ones (perhaps because prior arrests among White people signal relatively more involvement in crime than prior arrests among Black people). If a model is then fit to all the data, which would average the two trends, prior arrests for Black people would pre-

4 Human Rights Watch (2008).

dict a higher risk score than the truth, and prior arrests for White people would predict a lower risk score than the truth.

3 **Bias from differential censoring.** Predictive models are generally fit to only those defendants not detained while awaiting trial. Those detained, potentially because they were unable to make bail, are not included as there is no outcome related to new criminal activity or failure to appear (this is the "censorship"). This censorship can cause multiple problems. First, if the people who are detained are systematically different from the people who are not detained, the final models may not generalize: the models may not accurately predict risk for those people who were detained. Additionally, if detention patterns differ by racial group, bias may be introduced by fitting models using different subsets of the racial groups.

4 **Bias introduced when selecting thresholds for risk categories.** Scores produced by risk assessments are frequently grouped into categories for making release recommendations. Even if the initial, raw scores are unbiased, bias may be introduced when they are grouped, if the distribution of risk scores within those categories differs by race. Take, for example, a risk assessment that produces raw scores of integers ranging from 1 to 6. If detention is recommended for scores of both 5 and 6, there could be bias if the proportion of 5s to 6s is different for Black defendants than it is for White defendants.

5 **Bias introduced by relying on factors that override risk-assessment scores.** Guidelines for pretrial detention decisions typically incorporate information other than results from a risk assessment (for example, some jurisdictions may automatically "bump up" the effective risk score of a defendant if the arrest involves a violent crime). Bias may be introduced if this other information is both not additionally predictive of an outcome (beyond the baseline data used) and occurs differentially across groups. For example, if a particular targeted charge is more common among Black defendants than among White defendants, but that charge does not predict failure to appear beyond the baseline prediction, then a "bump-up" based on that charge may introduce bias that does not exist in the underlying risk-assessment algorithm.

6 **Bias in implementation.** Bias can be introduced in the implementation of the case pretrial process. For example, one source of bias in the pretrial process is differential decision-making by judges (and others who make detention decisions): Judges (or other stakeholders) may impose harsher release conditions on one group than another, even when they receive the same release recommendation for similar cases and contexts. One purpose of systematic risk assessment is to ameliorate such biases, but this potential for bias in implementation underscores the need to assess the bias of the fully implemented system as well as the risk-assessment tools themselves.

ASSESSING BIAS

The first three sources of bias discussed above can introduce bias into the risk assessment itself. There are three important points to make about assessing bias in risk-assessment scores:

1. **The validity of the estimates of bias in risk scores depends on several assumptions.** First, it assumes that the outcome that a risk-assessment model predicts is not measured with bias. This assumption can be addressed by considering outcomes that are less subject to bias because, for example, the crimes on which they focus are less prone to differential policing or arrest. The validity of estimates of bias in the risk scores also assumes that the relationships between the risk factors and the outcome are the same for all races. This assumption can be investigated by fitting models separately within different racial groups and comparing the results with the model fit to all racial groups pooled together. Additionally, the validity of bias estimates assumes that the censoring of outcomes for detained defendants does not introduce bias due to correlations between detention patterns and race. There may not be a way to assess this last point fully, although one could potentially use differences in judges' behavior to do so at least partially.

2. **There are different ways to define bias, or the desired goal: fairness.** Different stakeholders may put more priority on different definitions of fairness. The literature on the bias/fairness of predictive models is vast and it lacks consensus on the best definitions or measures. (It also lacks consistency in the names assigned to different concepts and measures.) It is important for researchers to focus on the questions various stakeholders (the court, the public, the defendants) would want answered to assess fairness.

3. **The extent of bias may change over time.** Jurisdictions across the country are implementing comprehensive criminal justice reforms. These reforms can have substantial effects on how pretrial risk assessments are used, on the measures that go into the assessments, on the composition of defendants for whom the risk assessments are applied, and on shifts in supervision practices during the pretrial periods. To the extent the reforms have effects, investigations of bias in risk assessment that are done with data from before the effects are fully realized may misrepresent levels of bias later. Because there may be continual changes in the defendant population and the pretrial legislative context, bias (as well as predictive performance) should be monitored frequently.

To further develop (2), above, this paper next reviews some questions stakeholders may ask about fairness and bias. Each question illustrates a different notion or definition of fairness, and each question maps to a different way of measuring bias empirically. As above, for ease of presentation these questions focus on race, but the same questions can be asked for other demographic groups for which it is important to assess bias/fairness. As will be discussed below, it is not possible to satisfy all definitions of bias simultaneously. Also, some ways of defining and measuring fairness have limitations.

1 **For a particular risk score is there the same chance of failure (of failure to appear, new criminal activity, or other adverse outcomes) in all racial groups during the pretrial period (or during a fixed period)? In other words, do risk scores mean the same thing regardless of race?** The statistical concept that answers this question is usually referred to as calibration or predictive parity. As noted earlier, the terminology is not consistent in the literature and other terms have been used as well. In any case, here failure rates are computed for each risk score and subgroup of interest. Subgroup failure rates are then compared within risk scores. Note that this question about fairness focuses on the risk *scores* (for example, scores that might range from 1 to 6) and not defendants' risk *categories* (for example, high risk or not), and the categories are what help judges decide whether or not to detain a person. Moving from a risk score to a risk category can result in very different conclusions about whether or not a risk assessment is fair. Therefore, when gauging the bias of a risk assessment, it is important to look not only at bias in the assessment but also at bias in the recommendations for how to use the results of the assessment.

It is also important to note that if the outcome itself is measured with bias, then perceived differences between two subgroups with the same risk score could be due to that bias. For example, if Black people are more likely to be arrested for the same behavior (for example, drug possession) then this appraisal is corrupt.

2 **Among those who would not fail if released, is the percentage assigned a particular risk score the same across all racial groups?** This statistical concept has been referred to as predictive parity or error rate balance (among other terms). This definition, while desirable for complete fairness, introduces some challenges. To achieve this type of fairness, a risk-assessment tool would need to treat people differently on the basis of race, which may not be ethical or even constitutional. Moreover, optimizing a model for this form of fairness makes it mathematically impossible to also achieve predictive parity, the first form of fairness discussed above.[5]

3 **Do all racial groups have the same distribution of risk scores?** It is understandable that stakeholders would want to know the answer to this question. However, it is impossible to untangle the extent to which different distributions of risk scores are due to bias — for example, from different arrest rates due to differential policing strategies or racism — or reflect actually differential rates of prior criminal activity and behavior that in turn are due to socioeconomic factors correlated with race.

4 **Do the risk scores produced by a risk assessment discriminate between failures and non-failures equally well across all racial groups**? There are several measures that approach this form of fairness question, such as relative rates of predictive accuracy, false positive rates, false negative rates, F1 scores (which combine both false positive and false negative rates), and area under the curve (AUC) measures. Unfortunately, these measures also are susceptible to the ten-

5 See, for example, Corbett-Davies et al. (2017); Chouldechova (2017).

sions between (1) and (2), above, and furthermore can be hard to interpret if the true distribution of risk is different for the different groups. Some of these measures (in particular the AUC) do avoid requiring a specification of a decision threshold for categorizing low and high risk, which can avoid some biases introduced by establishing such thresholds. Not specifying thresholds, however, could conceal biases that would arise when the risk tool was used in practice, as these tools are linked to an eventual hard threshold of detention or not.

ADDRESSING BIAS

While it is not possible to reliably measure or eliminate all of the sources of bias, as discussed above, there are steps researchers can take to mitigate and minimize bias. These are as follows:

1. **Identify risk-assessment outcomes that are likely to be less biased.** Biased outcomes can affect both the construction of a risk-assessment tool and its evaluation. To mitigate this bias, researchers might measure how risk scores correspond to new serious criminal activity or to new criminal activity that results in a conviction. They might also focus on chronic failure to appear rather than a single failure, as the latter might be more indicative of willful behavior rather than other factors such as lack of access to transportation. Such changes could also include creating modified outcome measures such as "seriousness-adjusted new criminal activity" (motivated by Hawken and Kleiman), so that risk assessments predict quantities that are as societally relevant as possible.[6]

2. **Identify risk factors that are likely to be less biased.** While measures of criminal history are logically strong predictors of future criminal behavior, they could also plausibly be biased, as discussed above. However, researchers could identify criminal history measures that are less susceptible to differential policing, differential prosecution practices, or racism. For example, one might argue that more serious crime is less vulnerable to bias. Similarly, there might be less bias in conviction rates than in arrest rates. In general, researchers should consider identifying aspects of people's case histories that are believed to be measured similarly in the different groups of interest. For example, researchers might discount all minor drug possession cases and focus on convictions rather than arrests when generating lists of potential risk factors.[7] Less biased risk factors in modeling lead to less biased risk assessment. Researchers can incorporate machine learning to help identify these less biased risk factors and to optimize model selection and stepwise variable selection for fairness (according to a particular definition and corresponding measure, as discussed above). Such a machine learning process will be most successful when the measure of the outcome being predicted has minimal bias, however. Otherwise, as discussed above, the process will hide the bias in the model.

6 Hawken and Kleiman (2016).

7 Beck and Blumstein (2017).

3 **Screen for risk factors that have the same relationship to the outcome across different races.** Given that risk-assessment models cannot specifically account for race (that is, include race as a factor), such screening will be an important step. One way of screening is to fit initial models that include race by risk factor interactions to identify risk factors that predict outcomes differentially, and then exclude those risk factors from subsequent modeling. Researchers can compare the overall predictive accuracy of the models fit to the restricted set of factors with the model fit to all factors to determine the cost in accuracy.

4 **Capture differential relationships between detention and subgroups when imputing censored outcomes of detained defendants.** One of the approaches to address the censoring problem when validating pretrial risk assessments is to impute missing outcomes for detained or partly detained defendants. This imputation could include race and other characteristics not used for the final risk model. (As imputation is only used to build the model, these defendant characteristics will not, in the end, be used as risk factors by the final predictive model and thus the overall risk assessment.) As long as the imputation captures any differential relationships between detention and subgroups, the subsequent model-fitting process will not be as vulnerable to biases from censoring.

5 **Assess bias in the decision-making guidelines, in particular those focused on thresholds for making risk-category distinctions and on the factors used to override initial recommendations based on risk scores.** When validating a risk assessment, researchers should assess the bias (and accuracy) implications of the decision-making guidelines before they are implemented and inform jurisdictions about the relative levels of bias in the factors and thresholds they are considering. Policymakers would then have to weigh any revealed trade-offs as they finalize their guidelines.

CONCLUSION

Bias can enter into pretrial risk assessments in many ways — some of which are frequently mentioned in popular discourse and some of which may be acknowledged less often. Not all sources of these biases can be measured or addressed fully, but researchers can do a better job of systematically identifying the threats to a risk assessment and the steps that can be taken to mitigate these biases.

It is at this point well known that many stakeholders — including the courts, the general public, and defendants — may make a priority of different notions of fairness. But this is only part of the story. As this brief has illustrated, a full understanding of potential biases in pretrial risk assessment is extremely complicated, and touches on not just the initial tools, but also the decision-making frameworks that guide their use and their actual implementation. Even the assessment of whether a system is fair can be corrupted by biased measures, which will only increase the degree of confusion in public discourse.

MDRC highlights the importance of bringing as much evidence to bear as possible. This brief has promoted a thorough and *ongoing* validation of risk-assessment tools and processes. It has also highlighted the essential, additional need for bigger-picture evaluations of these tools and processes' impacts on pretrial justice systems. Rigorous evaluation should assess how actual decisions are affected and how defendants' actual release outcomes are affected as well. Ultimately, humans rather than algorithms still make the decisions about which defendants are released and under what conditions. Therefore, it is not sufficient to assess and address the bias of the algorithms (the risk-assessment tools) — although that is a critical step. For a full picture, both researchers and practitioners need to assess and address bias in the final decisions and the effects those decisions have on defendants. This assessment may find that even though many risk-assessment tools are imperfect when evaluated in isolation, they may be parts of larger reform efforts that are substantial improvements compared with the starting point.

For more information about MDRC's Center for Criminal Justice Research
- www.mdrc.org/project/mdrc-center-criminal-justice-research
- Cindy Redcross, director (cindy.redcross@mdrc.org)

For more information about MDRC's Center for Data Insights
- www.mdrc.org/project/mdrc-center-data-insights-improving-programs-and-systems-actionable-data-science
- Kristin Porter, deputy director (Kristin.porter@mdrc.org)

For more information about MDRC's pretrial justice research:
- www.mdrc.org/issue/focus/pretrial-justice

REFERENCES

Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks." *ProPublica*. Website: www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Beck, Allen, and Alfred Blumstein. 2017. "Racial Disproportionality in U.S. State Prisons: Accounting for the Effects of Racial and Ethnic Differences in Criminal Involvement, Arrests, Sentencing, and Time." *Journal of Quantitative Criminology* 34, 3. Available online at www.researchgate.net/publication/318171222_Racial_Disproportionality_in_US_State_Prisons_Accounting_for_the_Effects_of_Racial_and_Ethnic_Differences_in_Criminal_Involvement_Arrests_Sentencing_and_Time_Served.

Calabresi, Massimo. 2014. "Exclusive: Attorney General Eric Holder to Oppose Data-Driven Sentencing." *TIME*, July 31. Available online at http://time.com/3061893/holder -to-oppose-data-driven-sentencing.

Chouldechova, Alexandra. 2017. "Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments." *Big Data* 5, 2: 153-163. Available online at https://arxiv.org/abs/1703.00056.

Corbett-Davies, Sam, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq. 2017. "Algorithmic Decision Making and the Cost of Fairness." *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Available online at https://arxiv.org/pdf/1701.08230.pdf.

Golub, Chloe Anderson, Cindy Redcross, and Erin Jacobs Valentine. 2019. *Evaluation of Pretrial Justice System*

Reforms That Use the Public Safety Assessment: Effects of New Jersey's Criminal Justice Reform. New York: MDRC.

Hawken, Angela, and Mark A.R. Kleiman. 2016. "Graduated Reintegration: Smoothing the Transition from Prison to Community." Policy Proposal 2016-03. Washington, DC: Brookings Institution.

Human Rights Watch. 2008. *Targeting Blacks: Drug Law Enforcement and Race in the United States*. New York: Human Rights Watch.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. 2017. "Human Decisions and Machine Predictions." *Quarterly Journal of Economics* 133, 1: 237-293.

Klingele, Cecilia. 2015. "The Promises and Perils of Evidence-Based Corrections." *Notre Dame Law Review* 91, 2: 537-584.

Redcross, Cindy, and Britt Henderson. 2019. *Evaluation of Pretrial Justice System Reforms That Use the Public Safety Assessment: Effects in Mecklenburg County, North Carolina*. New York: MDRC.

Stevenson, Meghan. 2019. "Assessing Risk Assessment in Action." *Minnesota Law Review* 103, 1: 303-384. Available online at https://minnesotalawreview.org/article/assessing-risk-assessment-in-action.

The authors thank Erin Valentine, Bret Barden, and Richard Hendra for their thoughtful contributions to the material in this research brief.

Dissemination of MDRC publications is supported by the following organizations and individuals that help finance MDRC's public policy outreach and expanding efforts to communicate the results and implications of our work to policymakers, practitioners, and others: The Annie E. Casey Foundation, Arnold Ventures, Charles and Lynn Schusterman Family Foundation, The Edna McConnell Clark Foundation, Ford Foundation, The George Gund Foundation, Daniel and Corinne Goldman, The Harry and Jeanette Weinberg Foundation, Inc., The JPB Foundation, The Joyce Foundation, The Kresge Foundation, and Sandler Foundation.

In addition, earnings from the MDRC Endowment help sustain our dissemination efforts. Contributors to the MDRC Endowment include Alcoa Foundation, The Ambrose Monell Foundation, Anheuser-Busch Foundation, Bristol-Myers Squibb Foundation, Charles Stewart Mott Foundation, Ford Foundation, The George Gund Foundation, The Grable Foundation, The Lizabeth and Frank Newman Charitable Foundation, The New York Times Company Foundation, Jan Nicholson, Paul H. O'Neill Charitable Foundation, John S. Reed, Sandler Foundation, and The Stupski Family Fund, as well as other individual contributors.

The findings and conclusions in this report do not necessarily represent the official positions or policies of the funders.

For information about MDRC and copies of our publications, see our website: www.mdrc.org.

NEW YORK
200 Vesey Street, 23rd Flr., New York, NY 10281
Tel: 212 532 3200

OAKLAND
475 14th Street, Suite 750, Oakland, CA 94612
Tel: 510 663 6372

WASHINGTON, DC
1990 M Street, NW, Suite 340
Washington, DC 20036

LOS ANGELES
11965 Venice Boulevard, Suite 402
Los Angeles, CA 90066