# Table of Contents

## Homelessness

**Stephen Eide**  
*Benchmarking Shelter Performance in New York: A Modest Proposal for Easing the City’s Homeless Crisis*  |  6

## Public Housing

**Howard Husock**  
*How New York’s Public Housing Fails the City’s New Poor*  |  19

## Transit

**Nicole Gelinas**  
*Not By Money Alone: Rethinking the MTA’s Infrastructure*  |  26

## Mental Illness

**D.J. Jaffe**  
*Our Insane Mental-Health System*  |  33

## Prisoner Reentry

**Howard Husock & Dean Ball**  
*How a Jobs-Driven Approach to Prisoner Reentry Can Reduce Recidivism*  |  36

## Education

**Marcus A. Winters**  
*Transforming Tenure: Using Value-Added Modeling to Identify Ineffective Teachers*  |  40

## Child Welfare

**Richard Gelles**  
*Creating an Effective Child Welfare System*  |  50

## Policing

**David Black**  
*Big Data on the Beat—Predictive Policing Has Arrived*  |  57
INTRODUCTION

By Michael Hendrix Director, State & Local Policy, MI

Despite the growing bill for public services—especially roads, schools, and public safety—government today is rarely efficient or effective or accountable. It is surely no coincidence that trust in government is declining. While roads and bridges in disrepair, failing schools, and dysfunctional criminal-justice systems are often laid at the feet of the federal government, state and local governments actually manage many of the public services that citizens use. City halls and state capitols often can have a greater impact on day-to-day life than Washington can.

The primary reason for failing public services is that government bureaucracies are not results-oriented. By and large, they concentrate on administrative process and not the material results of their efforts. If any of these policy failings are to be addressed, government must reorient itself. It must focus on serving its primary customer—the citizen. That, in turn, means paying for performance, making the most and best use of existing resources, and embracing innovation in essential services.

More than a quarter-century has passed since the publication of Reinventing Government, the bestselling book by Ted Gaebler and David Osborne. Their call for applying lessons from the private sector to the public sector sparked a newfound interest in more entrepreneurial and data-driven forms of management—particularly by measuring outcomes, sparking competition, and insisting on accountability.

The intervening years have shown their lessons to be helpful and desirable, but public leaders should not stop there. Metrics alone are no substitute for judgment, top-down accountability must be met with bottom-up buy-in, and competition is easily gamed.

To restore trust among citizens, the heads of state and local agencies need to build and own their goals from the bottom up (such as by empowering customer-facing public employees to identify problems and formulate solutions through public feedback), develop strategies to meet them, and measure results for accountability.

This booklet presents a playbook for achieving the efficient, effective, and accountable provision of social services, public safety, criminal justice, and transportation. The ideas are drawn largely from Manhattan Institute analyses of New York City, but they are applicable across the country. Some of the research highlights unique successes that New York has had in implementing results-oriented policy; many highlight failures.

Under New York mayor Bill de Blasio, the average stay in homeless shelters has surged from 293 days to 383 days in the past five years, even as spending has increased. The reason, Stephen Eide argues, is that the de Blasio administration eliminated performance benchmarks for shelters that had been established by his predecessor, Michael Bloomberg. Bloomberg’s program created an incentive for private shelters not simply to provide temporary housing with public dollars, but to find ways to permanently place single adults and families into their own housing. Absent the proper incentives, private, nonprofit service providers can be just as wasteful as the worst public bureaucracies; measuring their performance and rewarding only the good performers with more business is crucial.

The New York City Housing Authority (NYCHA) is the largest public-housing authority in the United States. NYCHA administers more than 16% of all public units in America, housing a population larger than Minneapolis. Yet Howard Husock finds that wait times for those seeking residence are long and turnover is low. Many of the city’s poorest find affordable public units to be effectively inaccessible. Those who do manage to find an apartment end up, more often than not, in dilapidated buildings. To achieve a more equitable outcome in the provision of public housing, policymakers should give priority to the
city’s underserved communities and ensure that the size of units correctly match individual and family needs. More broadly, city housing policy should “adopt policies that encourage a higher rate of turnover, particularly among tenants whose rising incomes place them firmly in the middle class.”

If public housing is a strain on local government services, public transit is where it buckles in New York City. The Metropolitan Transportation Authority’s (MTA) own data show it to be a model for cost overruns and schedule delays. Nicole Gelinas believes that the MTA should set its priorities straight: investing in the better buses and subways that the system’s ridership needs the most, rather than devoting outsize resources to commuter rail. So-called design-build contracts, which assign responsibility to a single contractor for designing and constructing a project, may help in out-of-control spending. Capturing the rise in real-estate values from this infrastructure investment could also help better finance the MTA’s continued upkeep and expansion, already burdened by rising debt-service and retiree-benefit costs.

The de Blasio administration spends $800 million a year on its ThriveNYC mental-health plan. But as D.J. Jaffe demonstrates, the plan prioritizes treating those with conditions like anxiety over those with severe mental illnesses, such as schizophrenia. Rather than being offered the treatment they need, individuals with serious mental illnesses are offloaded onto the city’s streets and eventually—due to the behavioral problems associated with untreated mental illness—its jails. This situation is needlessly expensive and inhumane; it also poses a significant threat to public safety. From the standpoint of public health and safety, government needs to focus on the specific areas where it can make the greatest difference: treating the most severely afflicted individuals.

Policing is one crucial area where New York City government shines. Through smart, focused policing, New York has experienced more than two decades of sustained, dramatic drops in crime. Today, the NYPD and other police departments are exploring new ways to reduce crime even further. Journalist and scholar David Black shows how they use Big Data to anticipate future crime and better allocate police resources. This strategy, known as predictive policing, takes officers off potentially unnecessary patrol routines and gives them time to engage in their communities more deeply—a critical and frequently overlooked aspect of law enforcement.

One vexing social problem, often overlooked, is how former prisoners can become law-abiding and functioning members of society. More than half of the offenders released from prisons are once again incarcerated within five years. State and local prisoner reentry programs are often of little help; their complex bureaucracies are rarely coordinated or held accountable. Howard Husock and Dean Ball draw on the experience of the Newark Prisoner Reentry Initiative, a partnership between the Manhattan Institute and Newark’s then-mayor, Cory Booker, to suggest that reentry policy should be focused predominantly on placing ex-offenders in a steady job. These reforms, proven by real-world experience in Newark, include eliminating regulatory barriers to employment and reorienting parole systems to focus on helping ex-offenders find work rather than on punishing them for minor violations of the rules.

Good public schools and teachers can dramatically improve the lives of their students. Yet it is difficult—often purposely so—to know which teachers are effective and which teachers are not. Marcus Winters believes that value-added modeling offers one way to see how much a single teacher contributes to a student’s progress over time. Experience has shown that periodically evaluating teacher performance throughout their careers offers them valuable feedback while helping schools identify ineffective teachers. School districts that have used value-added modeling found that it is a useful predictor of teacher performance over time. There is no single change that can magically reform failing schools and raise student achievement. But if New York City and other jurisdictions were to stop granting “automatic” tenure and start factoring teacher quality into their employment decisions, value-added modeling could help.

Protecting children from parental abuse and neglect is one crucial public service with a checkered record. Child welfare research has been slowly making progress in identifying evidence-based tools to help child welfare social workers, but government has not always been helpful. Richard Gelles notes the parade of blue-ribbon commissions and agency reorganizations that have done little to protect children but a lot to increase funding and staff. “The real problem dogging the U.S. child welfare system,” says Gelles, “is not insufficient funds but insufficient flexibility in how the existing funds may be used.”

As Americans pay more and more for public services, they naturally have come to expect better results. When better results are not forthcoming—and, indeed, when the services deteriorate—trust in government declines. Paying for performance and demanding accountability are a means to the end of serving citizens and fostering trust in community. These are the lessons of reinventing government that should animate city leaders in the 21st century.
This report looks at the place and function of shelter—temporary housing—within New York City’s homeless-services system. As homelessness has increased over time, so, too, has the strain placed on the city’s shelter system. New York now spends over $1 billion a year to provide temporary housing to 14,500 homeless single adults and 15,200 families, about 60,000 people in all (with thousands more living on the streets and in the subway system). And that does not count the hundreds of millions of dollars the city also spends on prevention efforts aimed at reducing the number entering shelters and rental-subsidy programs to facilitate their exit. Nonetheless, the “crisis,” as it is referred to by numerous observers and the city itself, shows little sign of abating.

In response to sustained criticism of his approach to the current crisis, Mayor Bill de Blasio has twice put forth plans to reform the delivery of homeless services. Improving shelters—making them safer and cleaner—has been his administration’s top priority. The administration has paid much less attention to the ability of shelter operators to move homeless adults and families out of their facilities and back into the community. It should do so and draw upon de Blasio’s predecessor’s approach.

Beginning in 2003, under Mayor Michael Bloomberg, New York City set up a system of performance benchmarking and financial incentives known as the Performance Incentive Program (PIP). This program measured and rewarded shelters’ effectiveness at reducing clients’ average length of stay and helping establish them in stable, independent housing. The Bloomberg administration viewed PIP as essential to its efforts in reducing the number of people living in shelters and their length of stay; there is evidence to suggest that PIP did just that. Almost every instance in which the city has managed to reduce the shelter census or average length of stay on an annual basis has occurred when PIP was active. After assuming office in January 2014, however, Mayor de Blasio let lapse the incentives and performance benchmarking for shelters.

The de Blasio administration should reinstitute a benchmarking and incentive program. Shelter operators should have their performance outcomes quantitatively evaluated and published in regularly issued reports, ranking them against peers that serve similar populations. Perfor-
performance should be mainly determined based on the rate at which shelter operators are placing adult- and family-shelter clients back into independent housing in the community, their average length of stay, and the rate at which formerly homeless people return to shelter. High performers should be rewarded with bonuses, and low performers should face a serious threat of loss of their contracts.

1 | A SHORT HISTORY OF NEW YORK CITY’S SHELTER SYSTEM

New York City’s shelter system developed in response to the emergence of the “modern” homelessness problem in the late 1970s. Prior to that time, the homeless population consisted mainly of indigent single men suffering from high rates of alcoholism.1 Today, most homeless New Yorkers are members of families, most of which are headed by single mothers (Figure 1).2 The homeless population now faces substance-abuse disorders other than alcoholism, as well as a high rate of serious mental illness.

New York’s response to modern homelessness evolved over four decades by five mayoral administrations. Two landmark events stand out.

The “Right” to Shelter

The first landmark event emerged from a lawsuit brought by advocates for the homeless against the city and state in 1979. Plaintiffs in Callahan v. Carey claimed that New York was abrogating its constitutional obligation to provide temporary housing for homeless single men, basing their argument for a “right to shelter” on a clause in the New York Constitution.3 Mayor Edward Koch settled the case in 1981 by signing a consent decree that conferred a right to shelter to single men (subsequent litigation and settlements extended the right to single women and families).4

The legal right to shelter may have reduced the percentage of unsheltered or street homelessness in New York City, though its effect is difficult to disentangle from other factors, such as New York’s long tradition of generosity in social-services spending and its harsh winter climate.5 Other cities that do not have a right to shelter (Philadelphia), or a much more qualified one (Boston and Washington, D.C.), have nearly the same rate of sheltered homeless (Figure 2). Counting unsheltered homeless is far from a precise science, so a difference of a few percentage points may not be consequential.

The right to shelter under the Callahan v. Carey consent decree has been interpreted to mean a right to imme-
diate shelter—adults and families must be placed in temporary housing the day they apply for it. Thus, city government has been forced to resort to forms of temporary housing—hotels and “cluster site” shelters jury-rigged in apartment buildings—that are widely seen as inferior to facilities designed for the specific purpose of housing the homeless. The result has been a very low quality of shelter offered to homeless families and single adults in recent years. More basically, the right to shelter has given legal advocates an outsize role in shaping the city’s response to homelessness. For many years, judges and lawyers from the Legal Aid Society and representing the Coalition for the Homeless have had more influence over how New York should respond to the homelessness challenge than many of the city’s elected representatives.6

**“Not-for-Profitization”**

The second landmark event in New York’s modern homeless policy was “The Way Home: A New Direction in Social Policy,” a 1992 report of the Cuomo Commission—a task force convened by Mayor David Dinkins, whose administration was facing a political crisis over its handling of homelessness.7 The report’s two most consequential recommendations were the creation of the Department of Homeless Services (DHS) as a separate city agency and a policy of privatization or, more precisely, “not-for-profitization.” The city was mostly to get out of the business of operating homeless shelters. Instead, the new DHS would set overall policy and “function as a ‘general contractor’ for the system,” with the frontline work of sheltering the homeless and transitioning them back into the community done by a network of government-funded nonprofits.8 Mayor Dinkins’s successor, Rudy Giuliani, implemented the Cuomo Commission’s recommendations.9 In the early 1990s, many governments were exploring ways to employ the private sector in pursuing public goods in areas such as transportation, elementary and secondary education, and sanitation services. Though the term “privatization” remains controversial in some circles, reliance on government-funded private contractors remains a foundational component of social policymaking in New York City. During FY 2017, the city administered more than 9,000 active contracts worth $22.6 billion for “human services.” About 400 of these contracts, worth $4.8 billion, were with the DHS.10

## 2 NEW YORK CITY’S SHELTER SYSTEM TODAY

New York City’s DHS oversees about 580 shelter facilities spread throughout the five boroughs.11 The city spent $1.3 billion on family- and adult-shelter services in FY 2017, a sum larger than outlays on libraries and parks combined.12 According to a 2015 report, New York City is host to roughly one-quarter of all the emergency-shelter beds in the nation.13 Homelessness in New York City has been termed a “crisis” by the administration, advocates, politicians of both parties, and various media organizations.14 As of late 2017, there were about 14,500 single adults and 15,200 families living in the DHS shelter system. Throughout de Blasio’s first term (January 2014 to December 2017), the number of families in shelters increased by 2,879, or 23%, while the number of single adults increased 4,657, or 47% (Figure 3).

![Figure 3: Trends in Single Adults and Families in Shelters, 2013-17](image3)

Providing temporary housing to single adults and families entails different managerial challenges. The single-adult population has higher rates of substance abuse and serious mental illness; families need larger units, privacy is a greater concern, and the city strives to place family clients in a shelter near their children’s schools.16 Single-adult homelessness, sheltered and unsheltered, is a significant driver of many so-called quality-of-life concerns in city neighborhoods, such as panhandling and public urination.17 Family and single-adult shelters are also funded differently. City tax revenues constitute a much larger proportion of the adult-shelter-services budget than that of family services (Figure 5).18 State regulations require that many services be provided to shelter cli-
HOMELESSNESS
Benchmarking Shelter Performance in New York: A Modest Proposal for Easing the City’s Homeless Crisis

exploration of housing possibilities with friends and family, the provision, directly or through a referral, of educational, employment, medical and mental-health services, and preparation for apartment viewings and interviews. Clients are instructed about how to comport themselves when viewing apartments and are even provided with professional attire, if necessary. Clients view apartments along with provider staff and are given help in moving when a suitable apartment is found. It is required to work are required to obtain and maintain employment.24 They must “apply for and use any benefits and resources that will reduce or eliminate the need for temporary housing assistance”25 and “actively seek housing other than temporary housing . . . and not unreasonably refuse or fail to accept any such housing.”26 Shelter clients must also “refrain from engaging in acts which endanger the health or safety of oneself or others, or which substantially and repeatedly interfere with the orderly operation of a temporary housing facility.”27 Noncompliance with DHS rules can lead to a client’s loss of shelter.28

Single Adults and Families

Since the 1990s, families seeking to enter the shelter system have been subjected to an eligibility process.29 All families must enter through the Prevention Assistance and Temporary Housing (PATH) intake center in the Bronx, where they receive a temporary shelter placement lasting, on average, 10 days, which turns into a permanent placement if they are found eligible. PATH staff scrutinizes a shelter applicant’s two-year housing history and interviews friends and family with whom they had recently lived to ascertain if they have nowhere else “safe and appropriate” to stay, even on a temporary basis, and are indeed in immediate need of emergency shelter.30

Intensive efforts are made to connect these families with prevention resources.31 Each month, DHS denies more than 1,000 petitions for family shelter. In his 2013 mayoral campaign, de Blasio criticized existing city shelter policy for its “unfair and overly punitive eligibility review rules that deny shelter to too many needy families.”32 Accordingly, in late 2015, the de Blasio administration sought greater leeway to grant shelter access from the state Office of Temporary and Disability Assistance (OTDA), which regulates city shelter services. However, one year later, after the eligibility rate had risen, the de Blasio administration asked the state to restore the previous authority of DHS to explore the possibility of keeping shelter applicants housed in an apartment leased to a family member or friend.33 Advocates opposed this move,34 which has brought eligibility rates closer to historical levels (Figure 6).

Single adults are not subject to an eligibility process.35 From a policy perspective, this is considered pru-
dent because of the risk of increasing street homelessness. New York’s most recent estimate of its unsheltered homeless population found that it was composed almost entirely of single adults. The Bloomberg administration tried to institute eligibility for singles, believing that it had taken sufficient steps to minimize the risk of expanding unsheltered homelessness. This action was strenuously opposed by homeless advocates. Ultimately, the courts denied the Bloomberg administration the ability to impose an eligibility process for single adults.

Progressives often cite increased use of government benefit resources as evidence of effective governance and express concern that some programs may be under-enrolled. Temporary housing benefits are an exception. In public debate, the shelter census tends to function as a scorecard that tracks whether the city is succeeding or failing at efforts to help the homeless. Shelter is seen somewhat like emergency-room usage: necessary but regrettable. This perception has fueled a growing emphasis on prevention and rental subsidies, designed to keep people from entering shelters and facilitating their exit, respectively. While spending on homeless services, in general, has doubled under the de Blasio administration, the growth has been driven mainly by programs other than shelter (Figure 7). Shelter now accounts for only 57% of all homeless-services spending.

Though more than 200,000 of New Yorkers have benefited from Mayor de Blasio’s prevention and rental-subsidy programs, the shelter census has continued to rise (Figure 3) and is expected to remain high for the foreseeable future. Mayor de Blasio projects a 4% decrease (2,500 people) in sheltered homelessness by 2022, which would leave levels still at a historical peak. Only a decrease in the number of homeless families is projected; the administration assumes that the number of homeless single adults and adult families (couples without children) will continue to increase in coming years.

The mayor has twice put forth plans to reform homeless services: the “90 Day Review,” released in April 2016, and the “Turning the Tide” plan of February 2017. In response to a series of spectacular tragedies and several highly critical reports, great emphasis has been placed on improving shelter conditions. Thousands of shelter inspections have been conducted and thousands of code violations addressed. The budget for shelter security has been doubled and now exceeds $200 million. The city government has brought in the NYPD to oversee safety operations in shelters. These increased investments in the physical plant of shelters and security have been accompanied by revisions to the city’s efforts at benchmarking levels of safety and cleanliness. A recent Daily News report criticized the city’s revisions to its definition of “critical incidents” on the grounds that it had exaggerated levels of safety in the shelter system. The city’s determination to improve shelter conditions has also driven its efforts to end contracts with a few grossly negligent providers, as well as its plan to phase out the use of cluster sites and hotels by 2021 and 2023, respectively. (In the near term, the planned phase-out of cluster sites will actually entail an increased use of hotels.) This is expected to enable the city to reduce the total number of shelter locations by almost 50%, even while it plans to expand “traditional” shelters by 90 new facilities and increase capacity at 30 existing ones.

3 SHELTER OUTCOMES VERSUS SHELTER CONDITIONS

In New York, shelter services have always meant more than just providing a clean, safe place to stay. In the most recent “Mayor’s Management Report,” DHS stipulates that one of its overarching goals is to “help individuals and families transition to permanent housing and self-sufficiency.” This entails that the agency not only ensure that “all temporary shelters for homeless individuals and families are clean, safe, and well-run” but also that it “[f]acilitate exits and minimize clients’ length of stay in shelters.”

Shelter outcomes—as opposed to shelter conditions—concern how effective providers are at moving clients toward an independent living situation in the community as quickly as possible.

The virtue of the not-for-profitization shelter system is not so much that private organizations would be better than the city at providing a clean, safe place to stay; instead, the theory was that these organizations would be more effective at establishing self-sufficiency.

Improving shelter outcomes and enhancing shelter conditions are two different tasks that sometimes are in tension with each other. The more comfortable an adult or a family feels in a temporary housing situation, the weaker the motivation could become to move back into the community. At a time when shelter conditions are
### Performance Incentive Program Result Card for Adult Homeless Services, 2009

<table>
<thead>
<tr>
<th>Shelter name</th>
<th>Shelter type</th>
<th>Percent of target achieved</th>
<th>Percent of census that are 9-month stayers</th>
<th>Percent of placed clients returning within 6 months</th>
<th>Percent of long-term stayers placed</th>
</tr>
</thead>
<tbody>
<tr>
<td>30th St. Assessment</td>
<td>Assessment*</td>
<td>157.41</td>
<td>N/A</td>
<td>5.94</td>
<td>N/A</td>
</tr>
<tr>
<td>Forbell Men</td>
<td>Assessment</td>
<td>151.50</td>
<td>15.14</td>
<td>4.10</td>
<td>57.89</td>
</tr>
<tr>
<td>HWC Assessment</td>
<td>Assessment</td>
<td>95.46</td>
<td>N/A</td>
<td>6.08</td>
<td>N/A</td>
</tr>
<tr>
<td>Franklin Assessment</td>
<td>Assessment</td>
<td>85.26</td>
<td>N/A</td>
<td>7.36</td>
<td>N/A</td>
</tr>
<tr>
<td>Atlantic M Assessment</td>
<td>Assessment</td>
<td>49.73</td>
<td>N/A</td>
<td>17.98</td>
<td>N/A</td>
</tr>
<tr>
<td>Santora</td>
<td>Employment</td>
<td>152.31</td>
<td>0.00</td>
<td>2.44</td>
<td>50.00</td>
</tr>
<tr>
<td>Palace Men</td>
<td>Employment</td>
<td>94.90</td>
<td>12.53</td>
<td>8.05</td>
<td>100.00</td>
</tr>
<tr>
<td>262 E. 3rd St.</td>
<td>Employment</td>
<td>90.38</td>
<td>10.02</td>
<td>17.14</td>
<td>54.29</td>
</tr>
<tr>
<td>Harlem</td>
<td>Employment</td>
<td>88.89</td>
<td>20.40</td>
<td>5.44</td>
<td>60.00</td>
</tr>
<tr>
<td>Hela SEC</td>
<td>Employment</td>
<td>88.64</td>
<td>6.60</td>
<td>15.82</td>
<td>33.33</td>
</tr>
<tr>
<td>Franklin ERI</td>
<td>General</td>
<td>158.14</td>
<td>6.36</td>
<td>12.35</td>
<td>N/A</td>
</tr>
<tr>
<td>85 Lexington</td>
<td>General</td>
<td>125.45</td>
<td>9.13</td>
<td>6.25</td>
<td>28.57</td>
</tr>
<tr>
<td>Broadway House</td>
<td>General</td>
<td>100.00</td>
<td>26.80</td>
<td>9.76</td>
<td>22.22</td>
</tr>
<tr>
<td>Borden Ave.</td>
<td>General</td>
<td>85.54</td>
<td>25.87</td>
<td>12.98</td>
<td>54.55</td>
</tr>
<tr>
<td>Porter Ave.</td>
<td>General</td>
<td>53.78</td>
<td>17.29</td>
<td>11.89</td>
<td>22.22</td>
</tr>
<tr>
<td>30th St.</td>
<td>General</td>
<td>37.42</td>
<td>13.94</td>
<td>8.33</td>
<td>9.68</td>
</tr>
<tr>
<td>Project Hospitality</td>
<td>Medical</td>
<td>106.25</td>
<td>0.00</td>
<td>0.00</td>
<td>60.00</td>
</tr>
<tr>
<td>Barler Free</td>
<td>Medical</td>
<td>104.17</td>
<td>39.07</td>
<td>2.63</td>
<td>53.85</td>
</tr>
<tr>
<td>Park Slope</td>
<td>Mental Health</td>
<td>191.67</td>
<td>11.21</td>
<td>4.85</td>
<td>36.29</td>
</tr>
<tr>
<td>George Daly</td>
<td>Mental Health</td>
<td>99.15</td>
<td>14.08</td>
<td>0.00</td>
<td>63.64</td>
</tr>
<tr>
<td>Valley Lodge</td>
<td>Mental Health</td>
<td>58.61</td>
<td>32.69</td>
<td>1.52</td>
<td>30.00</td>
</tr>
<tr>
<td>SCOW TLC</td>
<td>Mental Health</td>
<td>101.92</td>
<td>10.44</td>
<td>1.82</td>
<td>57.14</td>
</tr>
<tr>
<td>Park Ave.</td>
<td>Mental Health</td>
<td>66.67</td>
<td>43.27</td>
<td>1.96</td>
<td>30.91</td>
</tr>
<tr>
<td>Weston</td>
<td>Mental Health</td>
<td>104.17</td>
<td>14.11</td>
<td>5.88</td>
<td>N/A</td>
</tr>
<tr>
<td>New Haven</td>
<td>Mental Health</td>
<td>100.71</td>
<td>16.19</td>
<td>5.92</td>
<td>26.71</td>
</tr>
<tr>
<td>Ft. Washington</td>
<td>Mental Health</td>
<td>115.12</td>
<td>27.72</td>
<td>8.20</td>
<td>39.77</td>
</tr>
<tr>
<td>Renaissance</td>
<td>Mental Health</td>
<td>120.36</td>
<td>30.49</td>
<td>8.30</td>
<td>46.43</td>
</tr>
<tr>
<td>Susan’s Place</td>
<td>Mental Health</td>
<td>80.00</td>
<td>24.96</td>
<td>8.96</td>
<td>N/A</td>
</tr>
<tr>
<td>New Providence</td>
<td>Mental Health</td>
<td>94.64</td>
<td>18.81</td>
<td>9.26</td>
<td>66.67</td>
</tr>
<tr>
<td>St. Martin DePores</td>
<td>Mental Health</td>
<td>103.49</td>
<td>31.37</td>
<td>11.17</td>
<td>42.50</td>
</tr>
<tr>
<td>Schwartz</td>
<td>Mental Health</td>
<td>46.97</td>
<td>33.71</td>
<td>15.96</td>
<td>13.51</td>
</tr>
<tr>
<td>CAMBA Atlantic</td>
<td>Mental Health</td>
<td>114.53</td>
<td>29.48</td>
<td>16.67</td>
<td>N/A</td>
</tr>
<tr>
<td>HELP Women’s Center</td>
<td>Mental Health</td>
<td>78.65</td>
<td>19.13</td>
<td>18.46</td>
<td>9.09</td>
</tr>
<tr>
<td>Kingsboro MICA</td>
<td>Mental Health</td>
<td>104.35</td>
<td>29.44</td>
<td>24.59</td>
<td>30.29</td>
</tr>
<tr>
<td>Paloma Next Step</td>
<td>Next Step*</td>
<td>154.82</td>
<td>N/A</td>
<td>20.00</td>
<td>50.00</td>
</tr>
<tr>
<td>Willow Next Step</td>
<td>Next Step</td>
<td>116.07</td>
<td>N/A</td>
<td>10.97</td>
<td>28.57</td>
</tr>
<tr>
<td>Jamaica Next Step</td>
<td>Next Step</td>
<td>115.00</td>
<td>N/A</td>
<td>8.46</td>
<td>N/A</td>
</tr>
<tr>
<td>Schwartz Next Step</td>
<td>Next Step</td>
<td>42.50</td>
<td>N/A</td>
<td>16.43</td>
<td>N/A</td>
</tr>
<tr>
<td>Clarke Thomas Next Step</td>
<td>Next Step</td>
<td>42.25</td>
<td>N/A</td>
<td>15.89</td>
<td>N/A</td>
</tr>
<tr>
<td>CH Gay Assessment</td>
<td>Special Population</td>
<td>141.64</td>
<td>N/A</td>
<td>5.84</td>
<td>N/A</td>
</tr>
<tr>
<td>Turning Point</td>
<td>Special Population</td>
<td>96.00</td>
<td>8.31</td>
<td>9.43</td>
<td>N/A</td>
</tr>
<tr>
<td>Create Young Adult</td>
<td>Special Population</td>
<td>92.05</td>
<td>10.41</td>
<td>15.79</td>
<td>0.00</td>
</tr>
<tr>
<td>Bowery Mission</td>
<td>Substance Abuse</td>
<td>122.73</td>
<td>17.71</td>
<td>9.46</td>
<td>66.87</td>
</tr>
<tr>
<td>Palace Women</td>
<td>Substance Abuse</td>
<td>109.02</td>
<td>7.38</td>
<td>2.06</td>
<td>N/A</td>
</tr>
<tr>
<td>Kingsbom Star</td>
<td>Substance Abuse</td>
<td>106.32</td>
<td>26.88</td>
<td>24.19</td>
<td>32.79</td>
</tr>
<tr>
<td>Kanto</td>
<td>Substance Abuse</td>
<td>105.81</td>
<td>11.55</td>
<td>3.26</td>
<td>70.00</td>
</tr>
<tr>
<td>Casa Esperanza</td>
<td>Substance Abuse</td>
<td>104.38</td>
<td>21.43</td>
<td>2.94</td>
<td>100.00</td>
</tr>
<tr>
<td>81 E. 3rd St.</td>
<td>Substance Abuse</td>
<td>98.62</td>
<td>16.66</td>
<td>4.57</td>
<td>60.00</td>
</tr>
<tr>
<td>B. Kiefer</td>
<td>Substance Abuse</td>
<td>67.90</td>
<td>35.38</td>
<td>13.28</td>
<td>34.03</td>
</tr>
</tbody>
</table>

*An “assessment” shelter is where clients go after intake while it is being determined which shelter placement would best suit their needs; “next step” shelters were started by the Bloomberg administration. They provide intensive case management to long-term stayers.

Source: Obtained from a FOIL request.
said to be generally improving, like the present, it becomes all the more incumbent on shelter operators to work to place their clients in stable, independent housing.

**Bloomberg’s Performance Incentive Program**

Under Mayor de Blasio, however, the issue of shelter outcomes has been eclipsed by that of shelter conditions. The administration could have a policy rationale for ordering DHS’s priorities in this manner, or it could simply be responding to pressure from the many negative reports about shelter conditions and safety. Whatever the reason for de Blasio’s priorities, shelter outcomes received far more attention under his predecessor. In 2003, as noted, the Bloomberg administration launched the Performance Incentive Program (PIP), a system of benchmarking and financial incentives for shelter providers. Building on earlier efforts by the Giuliani administration, PIP measured providers’ rate of housing placements, clients’ length of stay, and the rate of return to shelter. The city set basic placement targets to ensure at least a minimum number of move-outs per year. Points were also awarded for “process”-type shelter functions, such as ensuring that clients were signed up for appropriate benefit programs and that their housing applications had been filed in a timely manner. Each shelter’s successes were tallied up to a final score. Shelter scores were published in quarterly rankings, where providers (and the city) could compare their own facilities against other providers that dealt with similar populations, as well as in an individualized report card. Providers were given financial rewards for performance and risked a loss of funding for failing to achieve their benchmarks. The financial incentives ranged around 10% of the facility’s base budget.

**Figures 8 and 9,** respectively, show examples of how Bloomberg’s PIP rated shelters: a systemwide “results card” for single adults from 2009; and a report card for an individual family shelter provider from 2012. Figure 8 shows that providers working with similar populations, the same housing market, and the same array of rental-subsidy programs placed at their disposal by the city met with varying results. Of the seven substance-abuse shelters, some experienced a 2% “recidivism” rate, i.e., a return to shelter within six months, whereas others’ rates were as high as 24%. Of the mental-health-shelter providers—which deal with some of the most challenging cases in the entire shelter system—most (11) experienced a recidivism rate of less than 10%. And for the cohort of 49 as a whole, 34 adult shelters in calendar 2009 met at least 90% of the target set by the city (indicating that standards in this particular year were not unrealistic), though some hit less than 50% of the target.

One year’s results should be kept in context: some shelters serve a small population, and their outcome data could be influenced, for instance, by a handful of difficult cases. But these results are striking and prompt important questions as to why some providers may be surpassing others in rates of placement and return to shelter.

The Bloomberg administration saw PIP as one part of a general approach to homelessness that included not only prevention and rental subsidies but specific obligations placed on clients as well as shelter providers to work toward the goal of self-sufficiency. From FY 2005 to FY 2010, the average length of stay for single adults declined every year. In 18 other instances, under Bloomberg, the aver-
age length of stay or average daily count declined for families or adults (Figure 10: years of decline have been shaded). Bloomberg cited PIP as one factor in these successes.58 His administration’s successes in keeping down the shelter census were, of course, only temporary; other factors, such as conditions in the low-rent housing market, would have to be explored.

Moreover, any comparison between the two mayors’ administrations has to consider that Bloomberg completed three terms in office, compared with de Blasio’s one. But it can be said that, since 2003, most instances in which the city managed to reduce the census or length of stay have occurred when PIP was active. The financial incentives for family-shelter providers.

The de Blasio administration has shut down providers for instances of egregious abuse.65 But scrutiny should also be applied to shelters that, though not abusing their clients, are not doing enough to move them along to self-sufficiency. The city is not doing enough to distinguish between poor-, mediocre-, and high-performing providers. As the de Blasio administration is embarking on its expansion of the shelter system and projecting high levels of homeless for years to come, this is an oversight that should be corrected promptly.

The way to do so is to relaunch the performance benchmarking program, both for family- and single-adult shelter providers. Providers with similar populations should be compared, in publicly available reports posted at least annually, in terms of their average length of stay, move-out rates, and rates of return to shelter. All such move-out metrics should constitute 75% of whatever score is given providers.

Shelter is an intervention that should be as brief and effective as possible. As with jails, mental hospitals, and foster care, there is a great risk that staying in a homeless shelter is preparing someone to be homeless, instead of for a life of independence. The public reports should provide more than rankings or tallies of final scores. They should indicate, in figures and language understandable to those outside the social-services world, exactly how well each of the city’s hundreds of shelter providers is succeeding at keeping its clients from staying too long and minimizing the rate of return after they’ve left.

What we measure depends on what we believe we want shelters to do. Some providers do important work in the area of employment services. And so long as there are thousands of seriously mentally ill people in the shelter system, mental-health services will be necessary, and we should be interested in which shelters do better work than others in this area. In the case of such special program shelters, success in mental health, substance-abuse treatment, and employment programs’ outcomes could also be tracked.

But none of those metrics should outweigh the average length of stay, rate of placements, and rates of return.

The public transparency aspect of performance benchmarking suffers when the metrics become overly individualized and complicated. Test scores, for example, are not the only way to evaluate whether a school is good, but they are an essential starting point. Concerns about overemphasis on short-term placements out of shelter can be met by monitoring rates of return: the goal is stable housing in the community, not “churn” from shelter to the community and then back to shelter. Special program shelters that excel at employment and treatment efforts are very likely to also experience low rates of return. (DHS has tracked system-wide rates of return for many years.) Providers should be free to continue to raise private funds to enhance their service offerings. But in terms of their core, government-funded mission, the emphasis should remain on sustainable move-outs.

In addition to benchmarking, the city administration should institute financial incentives. High performers should be rewarded with a financial bonus valued at 5%–10% of their base shelter contract. The ideal solution for low performers would be to put them
on watch and then, in the absence of improvement, transfer their work to a higher-performing provider with the capacity to take on the additional client load. The low-performing facility could remain in place and be taken over by another organization (i.e., changing shelter operators would not entail a controversy over where to site a new facility).

Benchmarking and financial incentives for shelter providers would be beneficial in the following ways:

• **Accountability is essential to the spirit of not-for-profitization.** As stated in the 1992 Cuomo Commission report: “Government should provide the performance-based incentives to encourage innovation and initiative among not-for-profit program operators. Results rather than process should be rewarded.”64 The current system grants independence to shelter providers but provides no incentives to encourage good results. DHS remains committed to client responsibility, but client responsibility is only meaningful in tandem with a rigorous system of provider accountability. Client responsibility and performance incentives for providers were seen by the Bloomberg administration as part of the same package of reforms aimed at “Minimiz[ing] the Duration of Homelessness.”65

• **Not all providers are equal.** Some social-services nonprofits that run shelters have been around for decades and take justifiable pride in their record. And all providers—on their websites, promotional materials, and appeals to donors—tout their achievements in improving clients’ lives. Only city government, through compiling and publishing standardized performance data, can give an objective accounting as to what degree these claims are true. The annual Mayor’s Management Report includes performance data for the shelter system as a whole but not on a per-provider basis. A provider that has managed, within two years, to work with five families to develop an exit strategy, navigate available benefits for them, and place them in stable housing, should be compensated more than another provider that manages only to work one family through the system during that same span. Yet under the current system, both receive the same compensation.

Properly designed, accountability should also have advantages for providers. Public benchmarking results would inform them about when certain of their shelters are falling short of the results of their peers. It also could enable strong providers to make their claims about outstanding performance more credible to potential donors.

• **Accountability would maximize use of a fixed resource.** Despite its abundance of government-funded housing, New York City has a long record of failing to make the most efficient use of existing resources.66 The legally created right to shelter may have made shelter a practically limitless resource, but quality shelter is undoubtedly a fixed resource. Indeed, as the de Blasio administration has noted, shelter quality has a tendency to be inversely proportional to shelter quantity.67

As difficult as this is to imagine, given the recent crisis in shelter conditions, an overemphasis on shelter quality could lead to a moral hazard problem, if life in shelters becomes more attractive than life in the housing that clients came from or what they are being offered by provider staff. This is not to make any assertive argument about what has been called a “draw to shelter” but rather to emphasize the critical importance of keeping the shelter census at a manageable level. At times in the past, clients in some instances were reluctant to take suitable units offered to them.68 Since providers are funded based solely on their operational costs, there is a certain risk of them becoming overly comfortable with their existing client load.

### 5 Conclusion

Homelessness remains a daunting challenge. The de Blasio administration has shifted the city’s target from the Bloomberg-era goal, to “overcome” homelessness, to managing the problem.69 Increased investment in prevention and rental subsidies should make the work of shelter providers easier. That was certainly the view of the Bloomberg administration. DHS commissioner Steven Banks recently testified that, though the rental market remains tight, with the aid of the rental-subsidy programs that the de Blasio administration set up, “[w]e’re funding substantial numbers of apartments” for shelter clients.70

Banks’s confidence in the administration’s ability to make headway with the crisis is backed up, in a sense, by data that show that only 7.5% of the city’s homeless population (5,755) is classified by the U.S. Department of Housing and Urban Development as “chronic”—meaning an individual or a head of household who suffers from a disability or diagnosable disorder and has been homeless for a year or four times in the preceding three years.71 This rate is less than half the rate of the chronic homeless population for the nation as a whole (Figure 1). In FY 2017, more than 17,000 single adults and family units exited shelter, about 7,800 of which did so without the assistance of any subsidy program.72 While the low-rent rental market remains tight and shelter recidivism remains a concern, it appears that independent living in the community remains in reach for most shelter clients.
For recent discussion of why communities experience different rates of homelessness, see Kevin Corinth, “Ending Homelessness: More Housing or Fewer Shelters?” AEI Economic Policy Working Paper 2015-12, Nov. 4, 2015; Kevin Corinth and David Lucas, “On the Relationship Between Climate and Homelessness,” AEI Economics Working Paper 2017-05, July 20, 2017. Corinth, “Ending Homelessness,” pp. 18–19, finds that “emergency shelter is strongly and positively associated with overall homeless counts, but only weakly associated with unsheltered homeless counts in the short run, and not at all in the long run. Meanwhile, it is most strongly and positively associated with the nonchronically homeless and families, segments which are least likely to end up on the street. Thus, there is little evidence that the majority of the emergency shelter inventory prevents unsheltered homelessness.”


9 DHS still directly operates nine shelters: six for single adults, two for families with children, and one for adult families; see NYC, Mayor’s Office of Operations, “Shelter Repair Scorecard.”

10 Office of the New York City Comptroller, “Checkbook NYC,” Human Services: Active Expense Contracts, FY 2017. These contracts cover several years, which is why these spending figures don’t align with other figures in this report about annual spending on homeless services.

11 Data are as of December 2017. NYC, Mayor’s Office of Operations, “Shelter Repair Scorecard.” Other small shelter systems are operated by the Department of Housing Preservation and Development, Department of Youth & Community Development, and Human Resources Administration (for the HIV/AIDS population and domestic-violence survivors). NYC, Department of Homeless Services, “DHS Local Law 31 report for the Month of December 2017.”


NYCRR, Chapter II, § 352.35(c)(3).
27 Ibid., § 352.35(c)(4).
28 Ibid., § 352.35(c).
32 Bill de Blasio for Mayor, “One New York, Rising Together,” 2013, p. 30: “No family should get caught in bureaucratic red tape while trying to access a shelter when they lack alternative housing options. Bill de Blasio will reform unfair and overly punitive eligibility review rules that deny shelter to too many needy families.”
36 Five of the 3,936 individuals counted were categorized as members of “households with only children”—a cohort that includes “persons under age 18, including children in one-child households, adolescent parents and their children, adolescent siblings, or other household configurations composed only of children.” Otherwise, every other unsheltered homeless individual was a single adult over 18. See HUD 2017 Continuum of Care Homeless Assistance Programs, Homeless Populations and Subpopulations, “NY-600 New York City CoC, Points in Time Date: 2/7/2017.”
37 NYC, Department of Homeless Services, DHS Strategic Plan FY 2012-2014.
41 Prevention programs include Homebase, legal services for tenants fighting eviction proceedings, and “one shot” emergency rental-assistance grants. Examples of rental subsidies include the Living in Communities (LINC), City Family Eviction Prevention Supplement/Family Exit Plan Supplement (CityFEPS), the Special Exit and Prevention Supplement (SEPS) voucher programs, and the reinstatement of NYCHA and Section 8 preferences for shelter clients. See “Turning the Tide,” chaps. 2–3.
43 “Turning the Tide,” pp. iii, 3, 86.
44 NYC, Department of Homeless Services, “Review of Homeless Service Agencies and Programs,” Apr. 11, 2016; “Turning the Tide.”
52 “Turning the Tide,” p. iii and ch. 5.
54 Ibid.
56 Rosegrant, “Linda Gibbs.”

NYC, Department of Homeless Services, Stats & Reports, “DHS Performance Report Archive.”


The Bushwick Economic Development Corp., Housing Bridge, and We Always Care are shelter providers that lost their contracts under de Blasio. See New York City Council, Committee on General Welfare, “Hearing Transcript, Oversight: Post-90 Day Review,” Apr. 20, 2017, p. 60, and idem, “Hearing Transcript, Oversight: PATH to Permanency,” June 27, 2017, p. 78.


See NYC, Department of Homeless Services, “Review of Homeless Service Agencies and Programs,” Apr. 11, 2016, p. 4: “Because the number of clients in shelter has increased as a result of the exponential shelter system census growth from 2011 into 2014 presented in the chart above, it has become increasingly difficult for DHS to adequately oversee and monitor providers, ensure safe, clean and secure conditions, and provide necessary services to clients.”


Main, Homelessness in New York City, pp. 193–94.

“Testimony of Steven Banks,” June 27, 2017, p. 149.


Scarceland, rigid buildingrules, and athriving economy thatattracts new workers have made New York one oftheworld’s mostexpensive cities in which to rentorbuyahome.²

Low-income New Yorkers suffer most fromthelack ofaffordable market housing. Many cram into tiny apartments. Others seek accommodation in the city’s vast public-housing system—spanning176,066 apartments in326 developments³ and home to one in 14 New Yorkers.⁴

Here, too, low-income city dwellers areill-served. Average waitingtimes for a New York City Housing Authority (NYCHA) unit stretch to 7.5 years.⁵ Those who endurethewait face high crime,⁶ dilapidated buildings,⁷ and scant access to supermarkets and other stores.⁸ NYCHA does little to ensure that its scarce public resources arefairly shared. Despite a shortage of units, NYCHA residents can stay indefinitely. Many do: the average tenant has lived in public housing for more than 18 years.⁹ Nor do residents face pressure to leave when theyrise outof poverty. More than one in ten NYCHA households have incomes greater than the New York City median ($53,000).¹⁰ Low turnover is not the only problem. More than a quarter of NYCHA residents live in apartments so spacious that bedrooms outnumber occupants.¹¹ Meanwhile, some groups are sharplyunderrepresented. Asians, for instance, account for 11.1% of the city’s households in poverty but make up only 4.7% of NYCHA households.¹² To create more equitable public housing, NYCHA should, among other things, prioritize outreach to underserved communities, as well as adopt policies thatencourage a higher rate ofturnover, particularly amongtenants whose rising incomes place them firmly in the middle class.
NYCHA’s Racial Disparities Are a Long-Standing Problem

The New York City Housing Authority, the largest public-housing authority in the U.S., aims “to increase opportunities for low- and moderate-income New Yorkers by providing safe, affordable housing and facilitating access to social and community services.” NYCHA was founded and its first buildings were dedicated in 1935, two years before the 1937 Wagner-Steagall Act ushered in federal support for public housing.

Today, NYCHA administers more than 16% of all public-housing units in the U.S. Its Queensbridge development is the largest public-housing complex in North America. If NYCHA’s authorized residents constituted a city, it would be America’s 48th most populous, larger than St. Louis, New Orleans, or Tampa.

Despite its vast size, NYCHA does not serve a representative cross-section of New York City’s poor. For example, 27.2% of the city’s households in poverty are white, but only 4.8% of NYCHA households are white. Likewise, 11.1% of households in poverty are Asian, but only 4.7% of NYCHA households are Asian. In contrast, blacks account for 25.8% of the city’s households in poverty but occupy 45.2% of NYCHA units. Hispanics, who account for 33.8% of households in poverty, occupy 44.7% of NYCHA units.

Faced with New York’s high market rents and long waiting lists for public housing (the average NYCHA applicant must wait 90 months), many poor New Yorkers—especially Asians and non–Puerto Rican Hispanic immigrants—have little choice but to live in overcrowded private apartments (Figure 1). Meanwhile, 26% of NYCHA residents live in apartments so spacious that bedrooms outnumber occupants (see Appendix, Figure 9).

Historically, NYCHA has served some types of poor households more than others. For instance, in 1980, blacks represented 27.6% of New York City households in poverty but made up 51.1% of all NYCHA households. In the same year, whites represented 14.9% of NYCHA households (Figure 2).

In recent years, the underrepresentation of certain groups has not improved. As Figure 2 shows, Asians accounted for 11% of New York’s impoverished population in 2015—up from 2% in 1980—but made up fewer than 5% of NYCHA residents. Blacks and Hispanics are sharply overrepresented in public housing, accounting for 25.8% and 33.8% of impoverished households, respectively, but 45.2% and 44.7% of NYCHA households.

How NYCHA Fails to Adequately Serve All Groups of Low-Income New Yorkers

Low Turnover

NYCHA is much less successful than most public-housing programs at accepting new tenants and encouraging old ones to move out. According to the U.S. Department of Housing and Urban Development (HUD), the average U.S. public-housing resident has spent 10 years in public housing; in New York City, the figure is more than 18 years. In 2016, 13% of America’s public-housing tenants received their apartments in the previous year, compared with only 3% of NYCHA’s tenants. Similarly, the average U.S. public-housing resident waits about two years for his apartment, compared with 7.5 years for the average NYCHA tenant.

Inefficient Use of Space

A significant amount of NYCHA’s housing stock is underused. According to HUD, 16% of U.S. public-housing residents are “over-housed” (i.e., they live in apartments with more bedrooms than occupants), com-
pared with 26% of NYCHA residents, as previously noted. In other words, whereas nearly a quarter of privately housed foreign-born Asian households and nearly 30% of foreign-born non–Puerto Rican Hispanic households live in overcrowded private apartments (see Figure 1), a large minority of NYCHA residents live in apartments that could comfortably house more people.

**Many Nonpoor Residents**

Only households below the poverty line can become NYCHA residents. Yet NYCHA residents who later rise above the poverty line are not required to leave. Instead, they merely pay more rent.28 In 2014, 10% of NYCHA households earned more than the city median of $53,000 (Figure 3).

**Lack of Racial Diversity**

More than 90% of NYCHA households are black or Hispanic, though both groups together make up only 60% of New York households in poverty.29 Some NYCHA developments are even more narrowly segregated (see Appendix, Figures 5–8). For example, 12,977 households live in 31 developments (of which 25 are in Brooklyn) that are at least 75% black.30

**Figure 4** shows developments by size, location, and percentage of households that are black and Hispanic. The closer a development is to the diagonal black line, the more its population is exclusively black and Hispanic; developments that touch the black line have only black and Hispanic residents. The lack of racial diversity in NYCHA developments may discourage applications from underrepresented groups, especially immigrants, who may be willing to endure overcrowded private housing in order to have a sufficiently large cultural support network.

### 3 Conclusion

To make public housing in New York better reflect the demographic profile of New York’s poor population, NYCHA should take the following steps. First, bolster outreach efforts to underrepresented groups, such as poor whites, Asians, and non–Puerto Rican Hispanics. Second, to give all poor New Yorkers a better chance at securing public housing, encourage a higher rate of turnover by adopting the following rules:

1. **10-year time limits.** Given the huge backlog of applicants, new NYCHA residents should not be allowed to stay in public housing longer than the U.S. average for public housing.

2. **Fixed-rent, long-term leases.** Instead of penalizing NYCHA residents for earning more income, adopting fixed-rent, long-term leases would offer residents an incentive to improve their finances, as well as save for the time when their NYCHA residency concludes.

3. **Mandatory transfers and voluntary buyouts for over-housed residents.** Transfers would downsize residents to apartments of an appropriate size. Buyouts would pay tenants to move to private housing. Residents with higher incomes might receive smaller payments; residents with deep, long-established ties to their development might receive larger payments.

Finally, it is worth noting that a long-term tenure in public housing should be considered a mixed blessing, at best. Those of moderate means who choose to pay low public-housing rents are forgoing the chance to purchase property. Faced with growing concern about the comparatively lower wealth of blacks compared with whites,31 one cannot ignore the fact that public housing across the U.S. is disproportionately black (45%).32 It is time to reimagine public housing as a way station—one where households can save but also strive, and one that serves the full range of those in need.
### APPENDIX: NYCHA’S MOST HEAVILY SEGREGATED DEVELOPMENTS

#### FIGURE 5.

<table>
<thead>
<tr>
<th>Development</th>
<th>Location; Community District</th>
<th>% Asian Households</th>
<th>% Asian Households in Community District</th>
</tr>
</thead>
<tbody>
<tr>
<td>LaGuardia Addition</td>
<td>Two Bridges, Manhattan; CD 3</td>
<td>70.1%</td>
<td>28.8%</td>
</tr>
<tr>
<td>Leavitt Street–34th Avenue</td>
<td>Flushing, Queens; CD 7</td>
<td>59.0%</td>
<td>46.7%</td>
</tr>
<tr>
<td>Rutgers Houses</td>
<td>Two Bridges, Manhattan; CD 3</td>
<td>54.5%*</td>
<td>28.8%</td>
</tr>
<tr>
<td>Lower East Side III</td>
<td>East Village, Manhattan; CD 3</td>
<td>48.2%</td>
<td>28.8%</td>
</tr>
<tr>
<td>45 Allen Street</td>
<td>Chinatown, Manhattan; CD 3</td>
<td>47.1%</td>
<td>28.8%</td>
</tr>
</tbody>
</table>

Excludes NYCHA developments with fewer than 50 households. All figures consider the head of household’s race. Asians make up 11.8% of New York City’s heads of households. * = excludes Section 8 transition households.


#### FIGURE 6.

<table>
<thead>
<tr>
<th>Development</th>
<th>Location; Community District</th>
<th>% Black Households</th>
<th>% Black Households in Community District</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rutland Towers</td>
<td>East Flatbush, Brooklyn; CD 17</td>
<td>94.9%</td>
<td>89.4%</td>
</tr>
<tr>
<td>Randolph Houses</td>
<td>Central Harlem, Manhattan; CD 10</td>
<td>86.2%</td>
<td>59.9%</td>
</tr>
<tr>
<td>Brevoort Houses</td>
<td>Bedford-Stuyvesant, Brooklyn; CD 3</td>
<td>86.1%</td>
<td>59.9%</td>
</tr>
<tr>
<td>Tapscott Street Rehab</td>
<td>Brownsville, Brooklyn; CD 16</td>
<td>85.1%</td>
<td>76.4%</td>
</tr>
<tr>
<td>Park Rock Rehab</td>
<td>Crown Heights, Brooklyn; CD 8</td>
<td>83.6%</td>
<td>61.7%</td>
</tr>
</tbody>
</table>

Excludes NYCHA developments with fewer than 50 households. Community district figures consider only the head of household’s race. Blacks make up 21.8% of New York City’s heads of households.


#### FIGURE 7.

<table>
<thead>
<tr>
<th>Development</th>
<th>Location; Community District</th>
<th>% Hispanic Households</th>
<th>% Hispanic Households in Community District</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fort Washington Avenue Rehab</td>
<td>Washington Heights, Manhattan; CD 12</td>
<td>89.7%</td>
<td>60.4%</td>
</tr>
<tr>
<td>West Tremont Avenues–Sadwick Avenue Area</td>
<td>Morris Heights, Bronx; CD 5</td>
<td>84.5%</td>
<td>65.9%</td>
</tr>
<tr>
<td>Washington Heights Rehab (Groups 1 and 2)</td>
<td>Washington Heights, Manhattan; CD 12</td>
<td>81.2%</td>
<td>60.4%</td>
</tr>
<tr>
<td>Highbridge Rehabs (Anderson Avenue)</td>
<td>Highbridge, Bronx; CD 4</td>
<td>80.6%</td>
<td>60.5%</td>
</tr>
<tr>
<td>Twin Parks East (Site 9)</td>
<td>Tremont, Bronx; CD 6</td>
<td>79.3%</td>
<td>61.5%*</td>
</tr>
</tbody>
</table>

*Bronx Community Districts 3 and 6. Excludes NYCHA developments with fewer than 50 households. Community district figures consider only the head of household’s race. Hispanics make up 24.9% of New York City’s heads of households.

### FIGURE 8.

**Most Heavily White NYCHA Developments**

<table>
<thead>
<tr>
<th>Development</th>
<th>Location; Community District</th>
<th>% White Households</th>
<th>% White Households in Community District</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haber</td>
<td>Coney Island, Brooklyn; CD 13</td>
<td>72.8%</td>
<td>68.5%</td>
</tr>
<tr>
<td>New Lane Area</td>
<td>Shore Acres, Staten Island; CD 1</td>
<td>56.4%</td>
<td>44.8%</td>
</tr>
<tr>
<td>Williams Plaza</td>
<td>Williamsburg, Brooklyn; CD 1</td>
<td>50.7%*</td>
<td>64.3%</td>
</tr>
<tr>
<td>Surfside Gardens</td>
<td>Coney Island, Brooklyn; CD 13</td>
<td>47.5%</td>
<td>68.5%</td>
</tr>
<tr>
<td>Taylor Street–Wythe Avenue</td>
<td>Williamsburg, Brooklyn; CD 1</td>
<td>43.3%</td>
<td>64.3%</td>
</tr>
</tbody>
</table>

Excludes NYCHA developments with fewer than 50 households. Community district figures consider only the head of household’s race. Whites make up 39.2% of New York City’s heads of households. * = excludes Section 8 Transition households.

*Source: 2011–15 American Community Survey 5-Year Estimates and 2016 NYCHA Resident Data Book*

### FIGURE 9.

**NYCHA Households by Number of Occupants and Bedrooms, Percentage of Total**

<table>
<thead>
<tr>
<th>Occupants</th>
<th>0 bedrooms</th>
<th>1 bedroom</th>
<th>2 bedrooms</th>
<th>3 bedrooms</th>
<th>4 bedrooms</th>
<th>5 bedrooms</th>
<th>6 bedrooms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.7%</td>
<td>17.8%</td>
<td>10.8%</td>
<td>3.2%</td>
<td>0.2%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>2</td>
<td>0.4%</td>
<td>8.7%</td>
<td>15.2%</td>
<td>4.0%</td>
<td>0.7%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>3</td>
<td>0.0%</td>
<td>2.4%</td>
<td>10.7%</td>
<td>4.4%</td>
<td>0.6%</td>
<td>0.2%</td>
<td>0.0%</td>
</tr>
<tr>
<td>4</td>
<td>0.0%</td>
<td>0.8%</td>
<td>3.9%</td>
<td>4.2%</td>
<td>0.5%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>5</td>
<td>0.0%</td>
<td>0.2%</td>
<td>0.8%</td>
<td>3.3%</td>
<td>0.5%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>6</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.2%</td>
<td>1.0%</td>
<td>0.1%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>7</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.1%</td>
<td>0.6%</td>
<td>0.2%</td>
<td>0.0%</td>
</tr>
<tr>
<td>8</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.1%</td>
</tr>
<tr>
<td>9</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.1%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>10</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.1%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>11</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.1%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

*For example, 3.2% of NYCHA households have one occupant living in a three bedroom apartment. 2.4% of NYCHA households have three occupants living in a one bedroom apartment.*

*Source: “2014 Housing and Vacancy Survey,” U.S. Census Bureau*
Research assistance provided by Connor Harris.

1 All statistics derived from the New York City Housing and Vacancy Survey are the author’s calculations. Except where otherwise noted, all statistics derived from the U.S. Census, American Community Survey, and Current Population Survey are the author’s calculations from the University of Minnesota’s IPUMS-CPS and IPUMS-USA databases.


3 “NYCHA 2017 Fact Sheet,” New York City Housing Authority.

4 Ibid.


7 See, e.g., Barbara Huang, “NYCHA Residents Rally for Repairs, Funding,” Mott Haven Herald, Mar. 9, 2017.


9 “Assisted Housing.”

10 Author’s calculations and “2014 Data Tables: New York City Housing and Vacancy Survey,” U.S. Census Bureau.

11 Ibid.

12 See “NYCHA Resident Data Book Summary,” New York City Housing Authority; and “American Community Survey: 2015 Data Release New and Notable,” U.S. Census Bureau. In this paper, the labels “Asian,” “black,” and “white” exclude Hispanics.

13 NYCHA also offers “Section 8” rental vouchers that cover 86,194 units of private housing. See “NYCHA 2017 Fact Sheet,” New York City Housing Authority.

14 “About NYCHA,” New York City Housing Authority.

15 “NYCHA 2017 Fact Sheet.”


17 According to the U.S. Department of Housing and Urban Development, there are 1,074,437 public-housing units in the U.S., of which NYCHA manages 174,484. NYCHA itself gives an even higher figure, 176,066. See “NYCHA 2017 Fact Sheet.”

18 Ibid.

19 “City and Town Population Totals Tables: 2010–2016,” United States Census Bureau; and “NYCHA 2017 Fact Sheet.”

20 “NYCHA Resident Data Book Summary.”

21 “Assisted Housing.”

22 56% of NYCHA’s Hispanic households are of Puerto Rican ethnicity, though Puerto Ricans make up 34.2% of New York’s Hispanic households. See “2011–15 American Community Survey 5-Year Estimates,” U.S. Census Bureau; and “2014 Data Tables.”

23 This waste occurs despite a NYCHA rule that gives priority to applicants with more than two residents per bedroom. See “Tenant Selection and Assignment Plan,” New York City Housing Authority, Sept. 23, 2016.

24 “Assisted Housing.” NYCHA’s own figure is 21.4 years. We do not know the cause of the discrepancy. See “Special Tabulation of Resident Characteristics,” NYC.gov.

25 Ibid.

26 Ibid.

27 All NYCHA residents pay 30% of their income in rent.

28 “Poverty Tool: Data Tool,” NYC.gov.

29 “NYCHA Resident Data Book Summary.”


31 “Assisted Housing.”
TRANSIT
As New York’s deteriorating mass-transit system grabbed the attention of state and local leaders this past summer, elected officials nodded their heads in agreement that the state-controlled Metropolitan Transportation Authority (MTA) needs more money for capital improvements. The officials disagreed only on which level of government—the state or the city—should provide these new resources, and through what mechanism.

True enough, the MTA does need additional financial resources. But neither the state nor the city should provide these resources until the MTA can prove that it can perform better in choosing, executing, and financing capital projects for the benefit of New York City residents, visitors, and workers. For more than a decade, the MTA has failed to demonstrate that it can invest existing resources wisely. First, the authority has favored the improvement and expansion of commuter-rail lines over subways, even as ridership growth on subways has exceeded ridership growth on commuter-rail lines. Second, the MTA has failed to efficiently manage its large projects, with cost overruns and schedule delays overwhelming the downstate region’s ability to plan for the future. Finally, the MTA has failed to demonstrate success with new ways to deliver projects to save money and time, as well as new ways to finance projects to better reflect who benefits from them.
The MTA is responsible for New York City’s subway and bus systems as well as for the region’s two commuter-rail systems: Metro-North Railroad and the Long Island Rail Road (LIRR). The MTA is an independent nonprofit corporation, but New York’s governor appoints the largest share of MTA board members as well as the authority’s chairman.

To fulfill its responsibilities, the MTA has two budgets: a $15.6 billion annual operating budget for day-to-day expenses; and a five-year capital plan through which it invests in long-term physical assets. Since 1982, under successive capital plans, the MTA has invested nearly $160 billion in such assets. Over the full 35 years, the MTA has consistently replaced track and equipment, replaced and refurbished aging subway cars and railcars, rehabilitated stations and maintenance buildings, and upgraded select signal systems. Over the past 15 years, it has also embarked on expansion projects, including the Second Avenue Subway, along Manhattan’s Upper East Side; and the East Side Access project, to connect the LIRR to Grand Central Terminal.

The MTA’s current $32.5 billion capital plan covers the years 2015 through 2019. Under this plan, to which the city and state agreed in 2015, the federal government is expected to provide $7.6 billion in funding, the state of New York has committed $8.5 billion, and the city of New York will provide $2.5 billion. The MTA itself will raise $9.9 billion by issuing long-term bonds and $4 billion from “other MTA sources,” including real-estate sales. As MTA service slipped during 2017’s so-called summer of hell, the authority said that its capital plan was not sufficient and that the city of New York was not doing its fair share. In 2017, MTA officials said that the authority likely would require “additional capital investment of approximately $8 billion” to invest in a new signaling system, new subway cars, and “modern communications technology.” The MTA, backed by Governor Andrew Cuomo, wants the city to pay more for these upgrades. Neither the MTA nor the governor specified a figure.

New York City’s leadership agrees that residents should pay more, but differs on how, and it also requests that the MTA change its priorities. Mayor Bill de Blasio has proposed a millionaires’ tax that would require the approval of the governor; the governor, on the other hand, has said that he is considering a congestion-pricing plan. De Blasio’s administration has also suggested that the MTA “re-allocate resources from less critical investments” so that “New York City subway and bus riders get a larger share of resources already available to the MTA.”

This paper explores two questions: Does the MTA need more money, as the governor and the mayor agree that it does? And should the MTA undergo significant reforms before it receives such resources, as the mayor would like to see?

2 IS THE MTA INVESTING ENOUGH OVERALL?

One key question is whether the current capital plan is sufficient. That is, without regard to priorities, is the MTA investing enough money to keep up with the natural physical depreciation of its infrastructure over time and with population and ridership growth?

The answer is no, for several reasons. The MTA began its capital plans 35 years ago, facing an infrastructure deficit. It had made only minimal investments for more than two decades prior to 1982. The MTA has not fully closed that deficit over the ensuing decades. The MTA estimates that it should spend $27 billion to $32.5 billion over five years to continue to bring the transit system into a “state of good repair” as well as to replace assets that it started buying or building in the early 1980s but have since become obsolete. Yet the MTA will spend just $19.3 billion over five years on such repairs and replacements, a 29%-41% shortfall (Figure 1).

Yet the MTA cannot cut back on improvements and expansions in order to fund more repairs and replacements. Many of the MTA’s system improvements are modest. They do not result in more passenger capacity but are upgrades to obsolete payment and safety systems. During the current capital plan, for example, the MTA will spend more than $400 million to upgrade its fare cards and fare-collection system. Such upgrades will bring New York only up to the technology that other global cities, including London, Paris, and Hong Kong, have been using for more than a decade. Similarly, purchasing electric buses will not result in passenger improvements; rather, it is a gradual transition necessary to meet emissions targets.

Figure 1. The MTA’s $32 Billion in Planned Capital Investments, 2015-17 ($ millions)

- State of good repair
- Normal replacement
- System improvement
- Network expansion

Source Arthur’s calculations based on MTA Capital Program documents. Numbers do not add up to $32.5 billion due to rounding and the omission of cash management, administrative costs, and other miscellaneous projects.
upgrade to a superior technology. Examples of MTA system upgrades that are necessary but that do not result in superior service for the majority of passengers are listed in Figure 2.

The system improvements that do result in more passenger capacity are long overdue. For example, the MTA continues to modernize signals on the subways, upgrading from a mechanical to a digital system, allocating $2 billion for this purpose in the current capital plan. These upgrades allow the MTA to run more trains per hour, resulting in additional rush-hour passenger capacity. Yet over more than a decade, the MTA has completed such upgrades on only two of 25 subway lines: the L and the No. 7 (the latter will go into service in 2018). This work cannot be put off for another generation.

Moreover, the MTA’s major expansion projects—such as the Second Avenue Subway—have not kept pace with the MTA region’s population growth and ridership over nearly two decades (Figure 3).

The answer to the MTA’s current woes, then, is not to cut back on the financial resources that it devotes to capital investment. In fact, under the right conditions, the MTA should increase such investments.

3 IS THE MTA INVESTING IN THE RIGHT PROJECTS?

The answer to this question, unfortunately, is no. The MTA is investing its money in the wrong projects relative to ridership and population growth. The current five-year capital plan devotes 71% of its expansion budget to the region’s two commuter railroads (Figure 4), 10 times the commuter railroad’s 7% share of MTA ridership (Figure 5). Under this plan, the MTA will invest a further $2.4 billion in the $10.2 billion East Side Access project to connect LIRR trains to Grand Central Terminal rather than to Penn Station, as well as $2 billion in building a third track on the LIRR’s main line, for example.

Both are worthy projects. Yet in an era of scarce resources, ridership levels do not justify the vast imbalance among the MTA’s priorities. East Side Access will make commutes easier for 162,000 commuters. A smaller, yet significant, investment to upgrade the signals on the Nos. 4, 5, and 6 subway lines would improve reliability for the 160,000 people who take a subway from Grand Central Terminal each day. Accelerating the second phase of the Second Avenue Subway, too, to extend it farther south would take ridership pressure from the Nos. 4, 5, and 6 trains nearby; already, nearly 200,000 people are using the first three stations of the Second Avenue Subway each day. This disparity in expansion funding relative to ridership is not a recent phenomenon. Since the MTA entered its expansion era at the turn of the 21st century, it has undertaken five major projects: the first phase of the Second Avenue Subway, East Side Access, the No. 7 subway extension, and the Fulton Street and South Ferry subway stations (Figure 6). Though four of these projects are and were New York City initiatives, they consti-
tute just half of the five projects’ total funding. Moreover, the city itself paid directly for the No. 7 subway extension, and federal funds paid for South Ferry as well as for $847 million of the Fulton Street project, both post-9/11 recovery initiatives.

Nor do subways and buses get a disproportionate share of capital funding for repairs, replacements, and modest improvements. With 93% of the MTA’s ridership, subways and buses garner 75% of such funding (Figure 7).

These allocations are contrary not only to ridership figures but also to recent growth in ridership. Over two decades, subway ridership has grown 55%, while commuter-rail ridership is up 36% (Figure 8). Both figures justify investment in upgrades and expansion but also help demonstrate that subways required relatively more attention and dollars.

The MTA’s capital priorities also do not reflect its funding sources (Figure 9). The MTA funds its operating budget, including its annual debt-service costs for capital expenditures, with fare, toll, and tax revenues. It gets 76% of its fare funding from New York City subway and bus riders (as well as a small portion, $7 million annually, from Staten Island Railway riders). The MTA gets 61% of its tax funding from New York City sources.8 In addition, New York City taxpayers provide nearly half of New York State income and sales taxes, with which the state indirectly pays for its own contribution to the MTA capital plan as well as for similar transportation-investment projects outside the downstate region.9 New York City residents also constitute a majority of the drivers over MTA’s most lucrative toll bridges, making up 74% of drivers on the Verrazano-Narrows Bridge, 71%-73% on the Triborough Bridge, and 56% on the Bronx-Whitestone Bridge.10

These fare and tax dollars do not tell the whole story. Suburban workers use New York City subways as part of their commute, and New York City employers pay payroll taxes to the MTA on behalf of workers who have commuted from the suburbs. Still, such payments do not justify the vast disparity in capital funding. Nor do commuter-rail riders pay a higher percentage of upgrade and expansion costs through the portion of their fares that goes toward debt service. The subway and bus systems’ farebox-recovery ratio, or the percentage of annual costs covered by fare payments, as opposed to by tax and other subsidy payments, is higher than the LIRR’s, and not substantially lower than Metro-North’s (Figure 10).

The Case for Better Strategic Planning and Better Project Execution

Much of New York City’s subway system now suffers from a lack of capacity for existing riders, particularly at peak hours; the MTA attributed 42% of its 57,164 train delays in July 2017 to overcrowding.11 Capital investments to ease such overcrowding, including modern signal systems, are worthy and necessary but are also extremely expensive. The subway system will require cumulative billions of dollars in investment and years of work for still-marginal improvements in system capacity.

In the meantime, the MTA could ease its capacity issues at a much lower price with more strategic investment in bus service. The MTA could begin to provide a parallel service aboveground to its most busy subway routes, just as London has done over the past two decades. In London, bus ridership has grown 69% since 2000, even as ridership in New York, after increasing steadily, has fallen by 15% since its 2008 peak.12 In its current capital plan, the MTA plans to spend $169 million on 190 articulated buses.13 Increasing this investment by $80 million, along with smaller-scale investments in depot and maintenance capacity, could support continuous bus service on protected express lanes along Madison, Fifth, and Lexington Avenues. Designating express lanes on these three thoroughfares would supplement subway capacity on busy lines by up to 10% at a much lower cost than large-scale subway constructions and upgrades entail. (Such a capacity increase would require the cooperation of New York City, which manages the streets.)

...and Better Project Execution

Just as the MTA prioritizes its projects poorly, it manages design, engineering, and construction poorly, resulting in wasted money and time. The more than tripling of East Side
Access’s original budget, for example, has left less money and management resources for critical projects such as the Second Avenue Subway.

Cost overruns also plague projects that get far less attention. The MTA’s project to modernize the signals on the No. 7 subway line from Queens to Manhattan, for example, was initially projected to cost $266 million when the MTA began it in 2007. As the MTA finishes the project over the next several months, the final cost will be $394 million. The project is also more than a year past its original November 2016 completion date. Similarly, the MTA’s project to rebuild the Cortlandt Street subway station, closed since 9/11, is now projected to cost $182 million, compared with an initial $115 million estimate. The project likely won’t be ready until well more than a year after its initial February 2018 projected completion date.

Over the past several years, the MTA consistently missed the goals that it set for itself on starting and completing projects. As Figure 11 shows, the MTA regularly completes only 65%–75% of the construction projects that it plans to begin or complete in any given year, raising concerns about how effectively it could spend a new influx of resources.

5 FINANCIAL CLOUDS THREATEN FUTURE MTA CAPITAL INVESTMENTS

Under its state stewardship, the MTA has increasingly relied on debt to fund its capital investments, and the current capital plan will raise the MTA’s existing $37.9 billion debt burden by 26%. Meanwhile, the MTA has failed to manage its other liabilities, including future burdens for pensions and retiree health care. The MTA’s growing debt and retiree burden harms its flexibility to invest in physical infrastructure in the future. Should the MTA continue on its current trajectory, its liabilities will exceed its assets in the near future, even as its asset base, accounting for normal depreciation of physical infrastructure, has increased (Figure 12).

6 CONCLUSION

If New York City’s population and tax base continue to grow, the city and regional transit system also will need to grow. Yet before the state and city allocate more resources to the state-run MTA, they should change the way the MTA approaches its capital investments.

First, the MTA needs to better order its projects for the benefit of city residents, visitors, and workers. As it stands, the MTA prioritizes commuter-rail projects that have value but improve service for fewer people, dollar for dollar, than a more assertive capital plan for subways and buses would. Such a plan can give priority to projects that ease capacity constraints for a lesser cost, including reliable bus service along the most crowded subway routes.

Second, the MTA needs to better execute its capital projects. Aware of the problem, the MTA has begun to follow the lead of other entities such as the Port Authority of New York and New Jersey by experimenting with project-delivery methods such as “design-build,” under which a single contractor team bears responsibility both for designing and constructing a project. This method can reduce the problem of contractor coordination, a problem that has plagued many MTA projects, including East Side Access. On bus service, the MTA should consider contracting out service on its heavily subsidized express bus routes, using any savings to provide additional service to defray subway crowding on busier corridors.

Additionally, state lawmakers and the governor should consider ways to better match the financing of a project to the beneficiaries of that project. East Side Access, for example, can encourage denser construction of housing along the Long Island commuting corridor; a special MTA assessment on new property could capture some of that value to defray construction costs over time. Finally, the MTA must better control its long-term liabilities to ensure that any capital-investment dollars, whether provided by the state or the city, are not crowded out by growing debt-service and retiree-benefit costs.
Research assistance provided by Connor Harris.

1 Permanent Citizens Advisory Committee to the MTA, “The Road Back: A Historic Review of the MTA Capital Program,” May 2012; Metropolitan Transportation Authority, MTA Capital Construction Program, 2010-2014, July 2013. All figures throughout are adjusted for inflation.


8 “The Money for the MTA’s Subway Crisis Plan Is in Governor Cuomo’s Budget,” New York City Office of the Mayor, July 27, 2017. All figures are one-way trips.

9 Total Income and Tax Liability by Place of Residence, New York State Dept. of Taxation and Finance, 2014 data. Taxable Sales and Purchases by Geography and Industry Through February 2017, New York State Dept. of Taxation and Finance.

10 Charles Komanoff, transportation economist, Sept. 2017 (analysis requested by the author).


13 MTA, MTA Capital Program Dashboard.


MENTAL ILLNESS
One year ago Wednesday, Deborah Danner, a mentally ill 66-year-old woman, was fatally shot by NYPD Sgt. Hugh Barry while he was responding to a 911 call about a person screaming and acting erratically. It wasn’t the first time police had been called to aid Danner, who neighbors said did well when she was on medication. Those previous calls went off without incident.

This time Barry found Danner—who had written eloquently about the “curse” of having schizophrenia—naked and armed with scissors. She put the scissors down, but when she came toward him with a bat in hand, he felt threatened and shot her with his gun rather than a Taser.

Mental-health industry leaders marked the occasion by gathering on the steps of City Hall today to demand that the Police Department—which is already giving thousands more officers training in how to handle mentally ill individuals—further reform its practices and that the mayor set up a new commission to force them to make changes.

They’re putting pressure in the wrong place. We’d have fewer Deborah Danner-type tragedies if advocates focused their fire not on the cops but on the chronic failures of our horribly inadequate mental-health system.

The death of Danner was a tragedy for her and her family, as it was for Barry and his family. But the reality is that police step in only after the mental-health system fails. And it frequently fails. According to the NYPD inspector general, there were 157,000 911 calls for emotionally disturbed persons in 2016. A mayoral task force reported that the mentally ill now fill 38% of all city jail cells. The city Department of Health and Mental Hygiene admits 90,000 city residents with serious mental illness received zero treatment in 2015.

Those are examples of mental-health system failures, not police failures.

The mental-health advocates have shown little interest in taking steps to stop the mental-illness-to-criminal-justice pipeline. While researching my book, I discovered that nationwide the mental-health industry primarily focuses on improving mental wellness in the masses, rather than treating the most seriously mentally ill, like Danner.

That exact approach is echoed in New York City, where Mayor de Blasio says he is focusing more resources than ever on mental-health problems. But the highest-functioning individuals
go into programs and the noncompliant, psychotic, and delusional are offloaded to police.

Danner is the poster child for the failure of the mental-health system. She was intelligent and literate and had a good job before taking ill. Prior to the shooting, she had been a member of the city’s finest program for the seriously mentally ill, Fountain House.

But as Executive Director Kenn Dudek wrote, she had dropped out, and because de Blasio’s city government cut Fountain House’s budget by half a million dollars, Fountain House couldn’t send a case manager to her home to reengage her.

De Blasio wasn’t lacking money. He spends $800 million on his ThriveNYC mental-health plan. But that plan has never been focused on helping the most seriously mentally ill, as the growing numbers of mentally ill homeless sleeping on the streets make clear.

Another program that might have helped Danner is Kendra’s Law. It allows courts to order mentally ill persons with a history of multiple arrests or multiple hospitalizations—Danner had 10—to stay in one year of treatment while they live in the community. Had she been under a Kendra’s Law order, she would have had a case manager assigned to help her. That might have avoided the need to call the police.

Criminal-justice officials should resist the protestors’ calls to participate in yet another joint mental health-criminal justice task force. We had joint task forces in 2008 and 2014. In both cases, the mental-health members focused on improving criminal-justice procedures rather than their own.

Police and criminal-justice officials should make unilateral recommendations to the mayor along the lines of those proposed by former NYPD commissioner Bill Bratton at a Manhattan Institute forum. He argued that to help the seriously mentally ill, the mental-health system should increase the number of hospital beds, loosen civil commitment procedures, and more robustly implement Kendra’s Law.

In fact, the police unions could propose these reforms at their own rally on the steps of City Hall to honor NYPD Officers Miosotis Familia, Eder Loor, Rafael Ramos, Wenjian Liu, William Fair, Phillip White, Paul Tuozzolo, Emmanuel Kwo, and all the other officers who have been injured or killed by persons with serious mental illness that the mental-health system failed to treat.

A rally to improve the mental-health system would be a truly fitting honor for Deborah Danner.
PRISONER REENTRY
The U.S. prison and jail populations have been steadily declining for much of the past decade, at least partly because of a nationwide drop in crime. Approximately 6.9 million people were incarcerated in 2013, the lowest figure since 2003, according to the Bureau of Justice Statistics (BJS).

Recidivism rates have not fallen comparably. In fact, the national three-year recidivism rate has remained flat since the 1990s, with approximately two-thirds of those released rearrested within three years, according to BJS estimates. The five-year recidivism rate over that same time frame exceeds 75%. The overall U.S. prison population almost certainly would have fallen more substantially if recidivism rates were lower. To do so, state governments seeking to cut costs and reform the criminal-justice system would do well to examine and experiment with ways to bring these rates down.

Former prisoners seeking to reenter the community and live a normal life face a variety of legal, social, and economic barriers. These include the difficulty of enrolling in academic and vocational programs, attaining public benefits such as housing, and finding a job. Government policy often only confounds their challenges: the Council of State Governments Justice Center, for example, counts more than 40,000 legal and regulatory barriers at the state and federal level across the country, including some 1,300 in New York state alone. For example, in many states it is difficult or impossible for ex-offenders to obtain the certificates or licenses necessary for employment in professions ranging from truck driving to interior design. Furthermore, many individuals exit prison with strong disincentives to work. Unpaid child support, for example, can be a tremendous burden. More than half of federal and state prisoners are parents, and the average incarcerated parent with a child-support case leaves prison owing $20,000. Often, these individuals are pushed into off-the-books employment to avoid collections of these sums, called arrearages, or they avoid the labor market altogether.

Some of these barriers are a result of the inevitable stigma of a criminal conviction. Some, however, reflect poor government policy. In many states and localities, prisoner reentry systems are in need of serious reform. Often, they are a complex and disconnected bureaucratic apparatus with no centralized definitions and mission, a lack of accountability for outcomes, and a lack of coordination among agencies. Even promising ideas are often derailed by poor implementation, quality metrics based on inputs rather than outputs, and unclear objectives. Much of prisoner reentry policy is ineffective, at best, as evidenced by high rates of recidivism.
across the country; and a hindrance to reentry to society, at worst.

The reform proposal outlined in this paper is neither abstract nor hypothetical. Its core elements—making outcomes a priority, achieved by an emphasis on rapid attachment to work—were implemented by the Manhattan Institute in concert with the Office of Reentry in Newark, New Jersey, as a part of the Institute’s Newark Prisoner Reentry Initiative (NPRI).7 This initiative was a multiyear collaboration with the city of Newark and then—mayor Cory Booker to experiment with innovative approaches to lowering recidivism.8 NPRI served more than 2,000 individuals with a six-month job retention rate of over 70% and a one-year recidivism rate of approximately 20%—far below the New Jersey state average. What follows, then, is a proposal whose effectiveness is supported by real-world experience.

A Plan for Reform

NPRI can serve as a model for municipalities looking for ways to coordinate and improve existing reentry services and, importantly, to reorient state-based parole departments toward employment rather than rearrest. Too often, parolees are monitored with an eye toward ensuring that they do not violate the terms of parole—the penalty for which may be a return to incarceration. Indeed, parole violations have increased more than 600% as a reason for arrests since 1978, climbing to more than 152,000 rearrests in 2012—accounting for nearly one-quarter of all arrests in 1978, climbing to more than 152,000 rearrests in 2012—accounting for nearly one-quarter of all arrests in that year.9 An employment-focused, NPRI-style model would, in contrast, call on parole officers to seek to assist ex-offenders in starting new lives by steering them to social-services agencies with proven records of job placement. It is important to note that, although not all released ex-offenders are on parole, a great many (nearly two-thirds in 2012)10 are—and parole and probation departments are established parts of government in counties and states across the country. Changing that model—to focus on formal employment, reestablishment of family ties, and reintegration into healthy community life—can hold the possibility of long-term reduction of the U.S. rate of recidivism.

Currently, prisoner reentry services are offered through a variety of state and local agencies, none of which are clearly designated as the central, or lead, agency. The first step in reform is to establish a lead agency for handling all prisoner reentry affairs. This agency would be directly responsible for managing reentry planning, spending, and oversight. Crucially, the agency would also set a coherent goal for itself and all the other agencies and programs offering reentry services. The next step is to establish measurable standards by which to evaluate performance. For example, rather than measure the number of ex-offenders served by a particular agency, an outcome-oriented standard would look at how many of those served attained a job, how many were able to maintain that job over time, and, of course, how many were subsequently rearrested.

In Newark, the leading agency was a newly created Office of Reentry, housed in the Mayor’s Department of Economic and Housing Development. NPRI emphasized rapid attachment to work and supported job placement and retention through intensive case management and mentoring. The initiative was built on the notion that success derives from regularly gathering data about the effectiveness of service providers and establishing a management structure that allows it to take action in response to what it learns from such data. The city’s Office of Reentry—among other things—exercised direct oversight and set expectations for the service providers. The office established a clear definition of successful reentry: gaining and retaining employment and avoiding re-incarceration.

A paycheck provides individuals with a way to pay for necessities, as well as key intangibles: dignity and hope. Moreover, a focus on the specific, tangible goal of employment helps organize the range of services to which ex-offenders may be exposed—and is likely to benefit even those who do not immediately find a job. Results from NPRI, as well as studies by MDRC and Public/Private Ventures regarding a similar approach in New York City and elsewhere, provide empirical support for this model.11 Finding individuals unsubsidized permanent employment may be more challenging in difficult economic climates, but this model uses the means of unsubsidized employment, where possible, as a springboard for addressing a variety of other reentry issues, including housing, legal assistance, substance abuse, and health care, rather than seeking to resolve these issues prior to an individual seeking employment. Central to this approach is individually focused “case management,” which is designed to ensure that the particular situation (and progress) of each participant is considered by the agency or service provider responsible.

With this approach and an accompanying set of performance metrics, the Newark Office of Prisoner Reentry was able to use “strategic contracting”—withholding a portion of the overall contract amount until key benchmarks for job placement, retention, wages, and recidivism were met. Additional clients were steered toward those contractors, among a group of six, that showed the best results. The office also provided ongoing advice and support for agencies seeking to improve their performance.

In sum, reform based on the successful NPRI experience should follow its model: establish a lead agency in overall charge of prisoner reentry that operates a shared data-management system to measure performance metrics among the service providers—the key metrics being job placement, retention, and recidivism—and implement strategic contracting. Over time, the lead agency would use the data to understand which programs are best at providing reentry services. Similarly, parole officers would know the best-performing programs and direct individuals to those programs, meaning that over time, clients and resources would be directed only to high-performing agencies, programs, and service providers. Poor-performing programs would be identified and
helped to improve or, if necessary, terminated.

One final major hurdle faced by ex-prisoners is child-support arrearages, often crippling large. Because of the potential in many states for wages to be garnished, arrearages can be a strong disincentive to employment. Much of the money is owed to state governments in repayment for TANF funds that the state made to the custodial parent, rather than to the custodial parent directly.12

A government cannot forgive child support owed to a parent, but it can forgive arrearages owed to itself. Moreover, the revenue forgone is likely to be inconsequential. It is a futile effort to attempt to collect thousands of dollars from released prisoners, a fact reflected in the very low rates of collection from such individuals.13

State governments should consider strategies to ensure that child-support arrearages do not become a disincentive to work. One idea: arrearages owed to the state would be forgiven or suspended upon exiting prison, according to a graduated schedule—if, and only if, an individual meets certain job-retention benchmarks.

Many municipalities do not prioritize prisoner reentry services because most of the consequences of ineffective prisoner reentry services—namely, recidivism—are borne by the state, and not the municipality itself. Although NPRI essentially added new responsibilities (and hence, costs) for Newark, the program helped keep thousands of people from being rearrested, ultimately saving New Jersey millions in annual corrections expenditures.

Since the cost of reform would be a significant burden for many cash-strapped municipalities, prisoner reentry is best coordinated at the state level but administered at the local level, where parole offices and others involved in the process are more likely to have specialized, area-specific knowledge. Many municipalities administer prisoner reentry services funded through WEA, the Workforce Education Act, and similar federal programs, which requires that applicants have a GED—something many ex-offenders do not have. But municipalities often lack an incentive to assume the cost of administering high-quality, comprehensive prisoner reentry services, because the cost of re-incarcerating individuals is borne by the state.

Massachusetts, for example, spends approximately $45,000 per inmate annually; the cost of prisoner reentry services at the municipal level, then, pales in comparison with the potential savings at the state level. In Newark, NPRI spent, on average, a combined $3,500 per individual on case management and job placement services. However, if the proposal outlined above were to be initiated at the state level, the state could offer municipalities a share of the money saved due to reduced recidivism rates of individuals in that municipality. This would not only encourage municipalities to prioritize reentry services; it would also foster competition among municipalities to reach increasingly low recidivism rates each year.

Conclusion

The reforms outlined in this paper are neither a cure-all nor a quick and easy means to reduce budget expenditures and increase the effectiveness of programs. But they are a starting point for state and local governments that seek to reduce recidivism and re-incorporate formerly incarcerated individuals into their families and society.

ENDNOTES

2 Ibid.
4 See, e.g., Council of State Governments Justice Center.
7 See my testimony before the U.S. Senate Judiciary Committee, July 21, 2010.
8 While the NPRI was a successful pilot program, it has since been folded into the Newark Career Training Center, which handles a wide variety of career services not limited to prisoner reentry.
10 Ibid.
13 Ibid.
Public school teachers in the United States are famously difficult to dismiss. The reason is simple: after three years on the job, most receive tenure—after a brief and subjective evaluation process (typically, a classroom visit or two by an administrator or another teacher) in which few receive negative ratings. Once tenured, teachers are armored against efforts to remove them, and most do not face any serious reevaluation to ensure that their skills stay up to standard. With this traditional approach, tenured teachers sometimes lose their positions for insubordination, criminal conduct, gross neglect, or other reasons—but almost never for simply being bad at the job.

This state of affairs protects teachers (both good and bad) quite well but is clearly harmful to students. The effects of a poor teacher, research has shown, haunt pupils for years afterward. Being assigned to such a teacher reduces the amount that a student learns in school and is associated with lower earnings in adulthood (in part because having an inadequate teacher makes a child more likely to have an early pregnancy and less likely to go to college). An education system that protects bad teachers does a grave disservice to the children in its care.

In recent years, some school districts have experimented with changes in tenure rules. They seek the power to remove ineffective teachers and, in some jurisdictions, to reevaluate teachers throughout their careers.

A keystone of this reform movement is the replacement of subjective evaluation with quantifiable measures of each teacher’s effectiveness. The quantitative method is known as value-added modeling (VAM), a statistical analysis of student scores that seeks to identify how much an individual teacher contributes to a pupil’s progress over the years. The use of VAM in teacher evaluations is growing, but the method remains extremely controversial. Critics often claim that it does not and cannot measure actual teacher quality.

This paper addresses that claim. Part I analyzes data from Florida public schools to show that a VAM score in a teacher’s third year is a good predictor of that teacher’s success in his or her fifth year. Having established that VAM is a useful predictive tool, Part II of the paper addresses the most effective ways that VAM can be used in tenure reform.
VAM is not a perfect measure of teacher quality because, like any statistical test, it is subject to random measurement errors. So it should not be regarded as the “magic bullet” solution to the problem of evaluating teacher performance. However, the method is reliable enough to be part of a sensible policy of tenure reform—one that replaces “automatic” tenure with rigorous evaluation of new candidates and periodic reexamination of those who have already received tenure.

1 INTRODUCTION

Tenure and the Problem of Teacher Quality

Bad teachers substantially harm a child’s prospects. Studies have found that an ineffective teacher can cost pupils as much as a grade level’s worth of learning during a single school year.1 Further, bad teachers—those who do not make any measurable contribution to their students’ advancement—make students more likely to have an early pregnancy, reduce the chances that they will go to college, and have a negative impact years later on their pupils’ earnings as adults.2 A wide body of research has shown that even as teacher quality is a school’s most important driver of achievement, teacher quality varies a great deal from classroom to classroom in public schools.3 Since 2009, a few school districts around the nation have been experimenting with changes to business as usual, seeking ways to improve the quality of their teachers. Though these districts remain a small minority, the reform effort has gathered steam, especially in the past year. One of its most controversial suggestions is the redefinition—or even elimination—of tenure for public school teachers.

For years, teachers’ unions and their supporters have described tenure as a necessary bulwark against arbitrary or discriminatory termination, which was a common practice before the advent of modern employment law and labor standards. But the current tenure system protects bad teachers as well as good ones. (Very few tenured teachers are ever forced to leave the classroom.) We know that teachers vary in quality and that removing less competent teachers has the potential to improve students’ education. Therefore, we can be sure that pupils are ill-served by a system that ensures that bad teachers cannot be fired. As its defenders like to point out, tenure ensures only that teachers receive due process before they are terminated. However, in most school systems, the required due process is so burdensome—and has so small a chance of success—that in practice, poor performance is rarely a firing offense. To be rid of a teacher for poor performance, in most public school systems, an administrator must carefully document several proofs of incompetence over a sustained period of time—in the form of botched lesson plans, improper classroom development, and observed poor practice. These “proof points” are inherently subjective, and each is contestable in the hearing process. Meanwhile, measurements of actual outcomes—how much students have learned in the teacher’s classroom—are rarely considered. This is why poor classroom performance is so rarely cited as a reason for dismissal. For instance, competence was mentioned in only eight of the 45 cases in which tenured teachers were terminated in New York City in 2008 and 2009. And six of those eight included other charges such as insubordination or misconduct.4

One might argue that worthy teachers with good records have earned some protection against the effects of a personal crisis or a rough year in the classroom. Tenure, though, is not reserved for proven educators.
On the contrary, public school teachers are offered lifetime tenure very early in their careers—usually after three years—and the offer seldom has much to do with their performance. As of 2011, according to a review of tenure laws by the National Council on Teacher Quality (2011), only eight states require that performance of a teacher’s students be central to deciding whether to award a teacher tenure. That actually represents considerable progress, since in 2009 the NCTQ found that not a single state awarded tenure primarily based on effectiveness. Moreover, in most American public schools, that early-career tenure decision is often the only systematic examination of a teacher’s worth. Tenured teachers are rarely reexamined to ensure that their skills are maintained.

Why are measurements of effectiveness given so little weight in tenure processes? The simple answer is that, until recently, such measures did not exist. Tenure rules were written when performance was evaluated entirely on the basis of a classroom visit or two by an experienced observer. School systems simply lacked any objective measure of the teacher’s contribution to student learning. Today, better measuring tools exist, but the rules remain as written. When tenure is decided, nearly all the teachers in a typical school system receive a satisfactory or higher rating.\(^5\)

School systems need a better approach to tenure. Job protection, if it is to be offered at all, should be restricted to the best teachers. And policies should permit reevaluations of once-worthy teachers be protected long after their performance has faltered. Most important, tenure should be related to meaningful and objective measurements of teaching effectiveness.

On this last point, modern statistical tools present a promising avenue for reform. These measures, used in tandem with traditional subjective measures of teacher quality, could help administrators make better-informed decisions about which teachers should receive tenure and which should be denied it. Statistical evaluations can also be used to identify experienced teachers who are performing poorly, with an objectivity that reduces the risk of a teacher being persecuted by an administrator.

To those dissatisfied with the status quo, one technique in particular seems to offer a good basis for reform, and it has been implemented in many recent attempts to change tenure rules in order to improve teacher quality. It is the method known as value-added modeling (VAM). VAM uses a complex statistical procedure to determine each teacher’s independent contribution to improvement in his or her students’ test scores.

Many school systems across the nation have recently used, or are currently considering using, VAM assessments when making employment decisions. For instance, under new laws passed in Colorado in 2010, Tennessee in 2011, and just recently in New Jersey, teachers in those states will lose their tenure if they receive below-satisfactory performance ratings in two consecutive years. Those ratings are based, in part, on VAM.

Some worry that because VAM is an imperfect measure of classroom effectiveness, it will incorrectly deny tenure protections to some effective teachers—or even cause good teachers to lose their jobs. If so, VAM’s negative impact might cancel out its benefits and result in no net improvement in the quality of a school district’s teaching staff. After all, research shows that VAM is an imprecise measure of a teacher’s true performance.\(^6\)

For this report, I test the premise that a teacher’s VAM score can help predict his or her future performance. I use data from Florida to replicate recent analyses by two scholars, Dan Goldhaber and Michael Hansen (2010),\(^7\) who used data from North Carolina. Consistent with their research, my results show that pre-tenure VAM scores are significantly related to student test-score performance in the teacher’s classroom in later years. These results indicate that VAM often contains meaningful information about a teacher’s future effectiveness, which can usefully inform employment decisions.

Obviously, the potential effects of any VAM-based tenure-reform policy would depend upon its design. Accordingly, the second part of this report looks at the number and type of teachers who would have been removed from the classroom (“deselected”) rather than tenured under different sorts of VAM-based policies, had those policies been in place in Florida when the data were collected. These comparisons show that the effects of such policies on teacher quality will depend on the standard that a teacher must meet to receive a satisfactory rating and on whether a teacher can lose tenure after it has been granted. These design issues, though important, should not obscure the fundamental point: VAM-based tenure policies hold considerable promise for removing consistently ineffective teachers and thus improving teacher quality throughout the public school system.

Before considering the method and results from this report, it is worth emphasizing that though the analysis here focuses only on the influence of VAM on teacher tenure decisions, real-world policies will quite sensibly use VAM as only one measure of effectiveness when rating teachers. Therefore, this report has put VAM-based tenure policies to a hard test: by evaluating the effect of using VAM alone to identify and remove ineffective teachers, it has placed more reliance on VAM than a real district would. That the VAM approach passes this test is a striking indication of its usefulness.

It is important to recall that this analysis was created to test the ability of VAM to identify low-performing teachers under the structure of the current system. That is, the analysis assumes that teachers and school systems will not respond to the new rules by changing their other behaviors. This is unlikely to be the case in any real-world application of tenure reform. Instead, teachers could reasonably be expected to respond to a reformed tenure system in several ways. The reformed system might, for example, attract a different sort of candidate. Further, teachers could...
respond to the new possibilities—not receiving tenure or being removed from the classroom—in ways that are good for students (by increasing their effort level), or that have unpredictable effects (changing their teaching style), or that could have negative effects (emphasizing only testable material in the classroom).

Additional theoretical and empirical research is needed to map the real-world effects of incorporating VAM-based measures of teacher quality into employment decisions. However, understanding the ability of VAM to predict future performance and the type of teacher identified as ineffective by a VAM-based system is an essential first step.

Balancing the Needs of Teachers and Pupils

Though VAM is a powerful technique, it is undoubtedly an imperfect measure of a teacher’s effectiveness. VAM is limited partly because it considers student performance only as measured by standardized tests, which are themselves imperfect measures of student achievement and account for only part of what school systems ask teachers to do. But even as a measure of the teacher’s contribution to student test scores, VAM has potentially serious limitations.

Critics of VAM analysis rightly point out that, as a statistical tool, VAM must contend with measurement error—the inevitable fact that measurements of the same thing, taken at different times, will vary, and some of this variation will be essentially random.

VAM-based measures of teacher performance can be quite imprecise. When VAM is used to inform tenure decisions, it is likely that some average and even above-average teachers could be removed from the classroom because of a low VAM score caused by random variation in measurement over the years, rather than their own failures. The influence of measurement error can be mitigated by statistical adjustments and by incorporating multiple years of student performance when evaluating any particular teacher. But measurement error cannot be eliminated.

From the perspective of teachers (and their unions), the collateral damage of even a single teacher losing tenure from an inaccurately low VAM score is unacceptable. However, the issue is not as cut-and-dried from the perspective of the student. A tenure-reform policy based on VAM will be an improvement for students if it removes enough low-performing teachers to improve overall teacher quality in a school district. If student achievement is our most pressing concern, we need to consider the possible consequences of VAM-based policies on whole districts, even as we acknowledge the potential for error in individual cases.

No evaluation system creates a perfect measure of an employee’s productivity. VAM, then, should not be judged against a nonexistent ideal but rather evaluated for its potential to improve on the current system’s ability to predict future performance. In the analyses that follow, this was my goal: to assess whether a tenure policy based on VAM would tend to improve a school district’s overall teacher quality.

2 VAM IS A RELIABLE PREDICTOR OF FUTURE PERFORMANCE

Following Goldhaber and Hansen’s work from North Carolina, my primary analysis uses a simple value-added model to estimate a teacher’s contribution to student test scores during the first two years in the classroom. I then evaluate the relationship between this measure and the achievement of students in the teacher’s classroom during his or her fifth year. If the previous VAM measure of teacher quality is a significant predictor of the teacher’s later achievement, we can conclude that VAM provides reliable information about a teacher’s future performance.

The analyses use detailed data about Florida students’ performance on the state’s annual high-stakes math and reading exams, the Florida Comprehensive Assessment Test (FCAT) in the spring semesters from 2002 through 2009. Though individuals are not identified by name, the data set permits the analyst to follow the performance of each student over time. It also includes identifying variables for each teacher and a variable used to match students to teachers in classrooms.

My analyses only include students in the fourth and fifth grades. In later grades, students change teachers for each subject, making the assessment of teacher impact far more difficult. Further, testing in Florida begins in the third grade, and the analysis requires a baseline achievement score for the year before the study period. Therefore, grades before fourth are not available for this method.

I used student reading scores to create a simple value-added model by grade and year (a later check showed that results would be similar had I used math scores). The model accounted for the impact on test scores of such observed student characteristics as race/ethnicity, gender, and socioeconomic status (as measured by whether the children were eligible for free or reduced-priced lunches). After controlling for these and other variables, I was able to arrive at the estimated contribution of individual teachers to their students’ test scores. With a measure of teacher impact in place for each student, I could then look at the data at the teacher level to develop a rolling measure of each teacher’s quality over the years. As we have mentioned, most school systems offer tenure after three years in the classroom. Therefore, I calculated each teacher’s average VAM score during his or her first three years in the classroom.

Finally, I took the measure of each teacher’s average value-added score during his or her first three years back to the student-year data set. I used the VAM from those first three years
to help predict each teacher’s students’ achievement in the teacher’s fifth year (the 2007–08 school year). What I was looking for was a significant and meaningful relationship between pre-tenure VAM score and the performance of students in the teacher’s future classroom years later.

**Relationship Between Pre-Tenure VAM and Later Student Performance**

The results of the analysis are reported in Table 1. The first column reports the results from a regression analysis (a statistical method for showing the relationship among several variables) in which I mapped the relationship between student achievement and a teacher’s having a master’s degree. (The master’s is often used as a proxy for skill and commitment in current evaluation systems.) Consistent with previous research, I find no relationship between a teacher having a master’s degree and student outcomes.

The second column reports the results of a regression analyzing the relationship between the teacher’s average VAM score during the first three years in the classroom and the performance of that teacher’s students during his or her fifth year in the classroom. The result shows a statistically significant and substantial relationship between the teacher’s pre-tenure average VAM score and achievement in that teacher’s classroom several years later. The third column shows that a control for whether the teacher has a master’s degree has no meaningful influence on the finding.

Results reported in Table 1 demonstrate that the value-added assessment of the teacher’s effectiveness prior to the tenure decision is a significant predictor of the teacher’s later effectiveness. Thus, VAM measures early in a teacher’s career appear to be good predictors of how well a teacher will perform in the future. As mentioned, this result is consistent with the previous findings of Goldhaber and Hansen, who used data from North Carolina; it is important to note that data from another state’s school system, based on data from a different standardized test, show the same relationship between early-career VAM scores and later student success.

### 3 COMPARING THE EFFECTS OF DIFFERENT VAM-BASED POLICIES

Accepting that VAM can help predict future success for teachers, I turn to the next practical question for school districts: How should VAM be incorporated into tenure policy?

Policymakers must first consider the level of performance that a teacher has to meet to avoid an ineffective rating. This bar must not be set too low, or the VAM will have little impact on quality. For instance, a VAM-based policy that removes a large school district’s single worst teacher might have a substantial effect for the few students who would have been assigned to that teacher’s classroom but would have an infinitesimal effect on overall teacher quality throughout the school system.

A second issue to consider is whether a teacher who receives tenure under a reformed system would keep it going forward (as is currently the case) or whether teachers could be continually reviewed. If tenure continues to be decided in teachers’ third year on the job and they experience no further significant reviews, the impact of any quality-improvement effort will be limited to teachers at the start of their careers. This means that the policy might affect too few teachers and do nothing about older teachers whose effectiveness is fading.

Finally, policymakers must consider how to use multiple years of VAM scores to assign tenure or identify teachers for removal. The measurement error inherent in VAM analysis, along with other administrative issues, should lead school systems to use multiyear measures when making employment decisions. Policymakers could respond to this need by comparing teacher performance using the average VAM score over a multiyear period or, as districts in Colorado and Tennessee have already done, by removing teachers after they receive consecutive poor ratings.

Table 2 reports the number of students in Florida who were attached to teachers who would have been fired according to different versions of a VAM-based policy: first, one that removes a teacher who has received a poor rating based on the previous three years’ performance; second, a policy that removes teachers only after they have demonstrated below-standard performance during their first three years in the classroom; and third, a policy that removes teachers who perform below a particular standard during consecutive years.

The table shows that different versions of a tenure reform policy would benefit different numbers of students. As would be expected, policies that simply raise the VAM score considered acceptable will affect a greater number of teachers, and thus stu-
The effect of a tenure-reform policy on overall teacher quality in the school system depends both on the number and quality of teachers denied tenure under such a policy. Figures 1 through 5 compare the distribution of the 2008–09 VAM scores of teachers who would have been deselected at the end of the 2007–08 school year, according to these different systems, with those of teachers who would have avoided removal.

Though each figure represents a different policy, all show that teachers who would have been fired in 2008–09 were less effective than teachers who would have survived review. However, the figures illustrate that some teachers who were observed to be performing at or above the mean in 2008–09 would have been fired according to any version of tenure reform. The risk—of firing teachers whose later performance is above average—increases as the standard for failure is set higher. For example, a policy that removes teachers performing below the 25th percentile sets a higher standard than a policy that removes those scoring below the fifth percentile. But that policy is more likely to remove teachers whose later effectiveness would prove to be well above average.

The figures also enable us to compare the later performance of teachers who would have been deselected according to different policy styles. As was done in Table 2, we consider the quality of teachers deselected according to a policy that: a) removes any teacher whose average VAM score over a three-year period was below the Xth percentile; b) removes only entering fourth-year teachers whose average VAM score over their first three years was below the Xth percentile among all teachers; or c) removes any teacher with a VAM score below the Xth percentile among all teachers for consecutive years.

The figures illustrate that the most conservative policy design—that is, the policy least likely to remove teachers who later perform well in the classroom—removes those teachers who score below the Xth percentile during consecutive years. As Table 2 illustrates, this is the policy design that removes the smallest number of teachers. On the other hand, a policy that removes any teacher whose average VAM score over a three-year period is below the Xth percentile will tend to remove more teachers who would later demonstrate themselves to be effective, though even this policy will tend to remove more ineffective teachers than effective ones.

4 CONCLUSION

Like previous research found in North Carolina, my analysis of Florida data found that pre-tenure VAM scores often provide information about a teacher’s future quality. Thus, VAM analysis can help replace “automatic” tenure with employment decisions based on reliable evaluations. It can be part of tenure reform and thus can contribute to improving public education in the United States.

But which tenure-reform policies would make best use of this technique? I addressed this question by pinpointing the teachers in the Florida data who would have been removed from the classroom according to several different types of policies and performance standards. I found that any VAM-based policy would have removed teachers who, on average, performed worse than their peers later in their careers.

However, different versions of VAM-based policies proved to have different consequences. Specifically, certain versions increased the risk that effective teachers (as measured by VAM) would be removed. For example, a policy could target teachers for removal if they have two or more periods of consecutive poor performance. Alternately, the policy could simply score teachers on an average of their performance ratings for a given number of years. I found that the latter policy was more likely than the former to result in the removal of effective teachers (teachers who, despite a “bad patch” in the records, would prove to be effective later). Another way to increase this risk of “false positives,” I found, was to set the performance bar high. Such policies, applied to the Florida data, would also have resulted in the removal of teachers who would later demonstrate effective performance.

These results tell tenure reformers that they should consider the number and type of teachers likely to be denied tenure or removed from the classroom under their proposed policies. This will help them design policies that balance the interests of students in need of great teachers and the legitimate interests of teachers concerned that they will be inappropriately removed from the classroom because of a randomly low VAM score.
The need for well-designed policies should not obscure the finding that public schools can indeed use VAM to help identify teachers for tenure or removal. Instead, these results underscore the importance of blending VAM with sound policies. This report does not argue that VAM should be used in isolation to evaluate teachers for tenure or to make any other employment decisions. VAM, as we have seen, is subject to random measurement errors, and so must be combined with other methods of teacher evaluation.

The lesson of this report and of other research is that VAM can be a useful piece of a comprehensive evaluation system. Claims that it is unreliable should be rejected. VAM, when combined with other evaluation methods and well-designed policies, can and should be part of a reformed system that improves teacher quality and thus gives America’s public school pupils a better start in life.

## ENDNOTES

1. E.g., Hanushek (1992) finds that students assigned to a teacher whose students have results in the 75th percentile (i.e., whose scores are better than three-quarters of their fellow pupils) will test one year and a half ahead of where they started when the school year is over. Students with teachers in the 25th percentile, on the other hand, end up with scores that are only a half-year better than their starting point.
4. E-mail correspondence with the Department of Education.
7. The analyses use a rich student-level panel data set acquired from the Florida K-20 data warehouse.
8. Consistent with previous research, I adjust the teacher effects according to the empirical Bayes estimator.

## REFERENCES


Transforming Tenure: Using Value-Added Modeling to Identify Ineffective Teachers

Figure 1
Deselect if Average 3 yr VAM Below 5th Percentile

Figure 2
Deselect if Average 3 yr VAM Below 10th Percentile

Figure 3
Deselect if Average 3 yr VAM Below 25th Percentile

Figure 4
Deselect if Consecutive VAM Below 5th Percentile

Figure 5
Deselect if Consecutive VAM Below 5th Percentile

Figure 6
Deselect if Consecutive VAM Below 5th Percentile
Figure 7
Deselect if Average 3 yr VAM Below 5th Percentile - Only 4th Year

Figure 8
Deselect if Consecutive VAM Below 10th Percentile - Only 4th Year

Figure 9
Deselect if Consecutive VAM Below 25th Percentile - Only 4th Year
CHILD WELFARE
If the nation had deliberately designed a system that would frustrate the professionals who staff it, anger the public who finance it, and abandon the children who depend on it, it could not have done a better job than the present child welfare system.

The indictment that opens this paper could have been written anytime in the past few months or years in the United States. It could have been written in New York City, for example, after the death of Zymere Perkins in September 2016, or Jaden Jordan in November 2016, or Bianca Abdul in March 2017, or the grievous injury to Kadiha Marrow in April 2017.1

New York’s child welfare agency is not the only one that could be criticized for failing to protect children. In Los Angeles, four former L.A. County social workers are to stand trial for the 2013 death of eight-year-old Gabriel Fernandez. The four were supposed to protect the boy, who was in the care of his mother and her boyfriend. But Fernandez was found tortured to death—burned, shot with BB pellets, and doused in pepper spray.2

In my own city of Philadelphia, 17 individuals were convicted of, or pled guilty to, a range of charges, from third-degree murder to perjury, in the starvation death of 14-year-old Daniele Kelly.3 Among the 17 individuals were:

Dana Poindexter, Department of Human Services (DHS) intake worker convicted of child endangering, recklessly endangerment, and perjury. Sentenced to two and a half to five years in prison.

Laura Sommerer, DHS social worker: pled guilty to child endangerment. Sentenced to four years’ probation.

Julius Murray, caseworker for the social-services contractor MultiEthnic Behavioral Health: pled guilty to involuntary manslaughter, conspiracy, and child endangerment. Also convicted of health-care fraud. Sentenced to four to eight years in prison for manslaughter, conspiracy, and endangerment; sentenced to 11 years in prison for health-care fraud.

Mickal Kamuvaka, MultiEthnic Behavioral Health CEO; convicted of involuntary manslaughter, child endangerment, perjury, criminal conspiracy, and forgery. Sentenced to 17.5 years in prison.4

The inability to protect endangered children is not limited to large urban centers. In Rhode Island, the state Department of Children, Youth and Families admitted that, in April 2016, nearly two thirds (63%) of its kinship placement homes were unlicensed. Unlicensed homes can put
Let me begin with my most controversial statement: we do not really want to truly protect children. It is based on a number of key points. First, parental rights have priority in child and family jurisprudence. A series of Supreme Court decisions—from Smith v. Organization of Foster Families for Equality and Reform, 431 U.S. 816 (1977), to Santosky v. Kramer, 455 U.S. 745 (1982)—limits the state’s ability to intervene in the raising of children by their parents, and it sets a high bar for states that wish to terminate parental rights. Nor is this wrong. Upholding parents’ liberty interest to raise their children without unwarranted government interference is appropriate, as is a high bar for terminating parental rights.

Federal law also bolsters parental rights. The Adoption Assistance and Child Welfare Act of 1980 primarily provides states with funds for out-of-home placement but includes the requirement that states make “reasonable efforts” to keep children with their birth parents or safely reunify children from out-of-home placement prior to seeking to terminate parental rights.

The actual functioning of the child protective service system illustrates
how low the likelihood is that a child who is suspected of being the victim of maltreatment will actually be placed in foster care (Figure 1). Despite anecdotal critiques that child protective service agencies are too quick to remove children from homes,8 the data indicate that the vast majority of reports and the majority of substantiated reports of child maltreatment do not result in the removal of a child from his or her parents.

Last and perhaps most important, the culture of the American child welfare system sees parents as the clients and family preservation as the core goal of child protective services. Law professor Elizabeth Bartholet, in two key publications,9 summarizes how, for the past three decades, advocates, policymakers, foundations, and agency administrators have privileged supporting and preserving parents over the safety of children. To be fair, those on the side of family preservation hold fast to the value that children do best when raised by their birth parents and close family members. And to bolster the value of family preservation, members of what Bartholet refers to as the “racial disproportionality movement”10 use data on the race of children removed from their families as a club to try to limit such removals.

The second approach of the family-preservation advocates is to continue rolling out ever new interventions, such as Intensive Family Preservation Services, Family Group Conferencing, and Alternative Response with claims that such interventions can both preserve families and ensure the safety of children. By the time there are data to disprove the claim of effectiveness of a family preservation intervention, a new intervention is rolled out with the same claims.11 A final mainstay of the effort to preserve families is the claim that, with adequate resources and if done properly, child welfare agencies can preserve families as well as ensure the safety of children.

Anecdotal evidence, such as the cases that open this paper, disproves the claim that families can be preserved and children kept safe. The claim violates the laws of probability theory: it is impossible to both reduce false positives (concluding that a child is at risk of abuse when the child is not) and false negatives (concluding that the child is safe when the child is at risk). Choosing the parent as client can significantly disadvantage the safety of the child. Choosing the child as client reduces parental rights. Child welfare agencies must choose the errors that they are willing to tolerate.

By focusing on the parent as the client of the child welfare system and privileging parents’ rights, child welfare systems in practice hold children’s development hostage while waiting and hoping that parents will engage in services and that the services will be effective. This system chooses not to ensure the safety and well-being of children in harm’s way.

2 WE DO NOT KNOW HOW TO PROTECT CHILDREN

Undoubtedly, many will push back strongly, even in anger, against my claim that the American child welfare system as a whole does not want to protect children. My second argument is that child welfare systems do not know how to protect children.

Decision Making

While billions of dollars are spent on supporting children in foster care and services to assist parents—including parenting classes and drug-treatment programs—the most important component and task of the child welfare system in the United States is decision making. I envision the child welfare system as a series of nine gates that begins with the decision to report suspected child abuse and ends with the decision to close the case—through a reunification (the most common outcome) or by termination of parental rights.

No matter how skilled and experienced the decision maker, the actual tools that are available for decision making are not remotely up to the task. In the vast majority of cases and in the vast majority of decisions—decisions as to whether to report a case of suspected maltreatment, whether to substantiate the report, whether to remove the child from the home, and how to close the case—the main tool is clinical judgment. What we know about clinical judgment is that its accuracy is no better than chance, and it introduces bias, such as racism and classism, into the decision-making process.12
While there have been some modest advancements in developing actuarial tools for assessing safety and risk and to inform decisions, the child welfare field continues to be reluctant to replace clinical judgment with any of them.

The most recent development in decision making is predictive analytics, or “big data.” Predictive analytics holds much greater promise of improving child welfare decision making than do clinical judgment, consensus risk assessment, and older forms of actuarial risk assessment such as structured decision making. But the child welfare field is slow to embrace this tool. The major concern is profiling: critics worry about minorities and poor families being unfairly profiled by statistical tools—although, of course, such families are already profiled by clinical judgment. Since predictive analytics validates the algorithms with actual data, initial biases will be factored out over time. Nonetheless, the child welfare field, with few exceptions (Allegheny County, Pennsylvania, and Los Angeles County), seems to prefer the bias of clinical judgment to the potential of predictive analytics.

**Prevention and Intervention**

Currently, the best-tested and validated tool available to the child welfare system for the prevention of child maltreatment is the Nurse-Family Partnership, which involves trained nurses making home visits to low-income mothers who have no previous live births. The visiting nurses have three goals: (1) to improve the outcome of the pregnancy by helping women with prenatal health; (2) to improve the child’s health and development by helping parents provide more sensitive and competent child care; and (3) to improve the parental life course by helping parents plan future pregnancies. David Olds and his colleagues have spent nearly three decades evaluating the effectiveness of the Nurse-Family Partnership program, including three separate random clinical trials with different populations. The program has positive effects: fewer childhood injuries and ingestions that may be associated with child abuse; and fewer substantiated reports of child maltreatment by participating parents.

The child welfare system has developed an extensive menu of interventions. Almost every case file I have reviewed requires parents to attend parenting classes. As a large proportion of caregivers who become involved in the child welfare system have substance abuse issues, substance abuse treatment is a standard intervention. Intensive Family Preservation Services, Family Group Conferencing, and Alternative Response (all mentioned in the previous section) are common interventions. Their singular problem is the lack of empirical evidence meeting the normal standards of scientific evidence that these interventions reduce the risk of child maltreatment and keep children safe.

While there are many documents about evidence-based practice in child welfare, very few interventions are truly evidence based. Among the most widely discussed, evaluated, and effective interventions are Multi-Systemic Therapy (MST) and Triple P (Positive Parenting Program).

One takeaway from a review of evaluated as well as unevaluated interventions and prevention programs is that the focus is primarily on the impact of the intervention on the parent or caregiver. Few of the interventions are designed for or tested for the impact of the evaluation on the safety and well-being of children. Again, the parent-as-client bias pervades the development of tools for intervention and prevention (with the nearly unique exception of Nurse-Family Partnerships). The implicit assumption for the interventions and evaluations is that if an evaluation allows a child to remain with his or her birth parents, the intervention is a success. Safety and well-being, and even achieving developmental potential, become subordinate goals.

Over the past decade, the child welfare field has endeavored to develop effective, evidence-based practices. Progress is slow, as would be expected, given the time it takes to develop, test, and replicate random clinical trials. But it is still fair to say that an evidence-based toolbox for child welfare practitioners is relatively sparse.

The scarcity of good tools is partially due to the time it takes to develop them. But it is also related to the resources available for development and testing.

**The “Insufficient Funding” Red Herring**

Without question, the first and most consistent “suspect” rounded up to explain a child welfare agency’s or system’s inability to protect children is lack of funding. Although numerous funding streams flow into the child welfare system, including Medicaid, Temporary Assistance to Needy Families, and Social Service Block Grants, the most substantial funding streams are Title IV-B and Title IV-E of the (amended) Social Security Act of 1935. The Family Preservation and Support Program was added to Title IV-B in 1993. Now called “Promoting Safe and Stable Families,” this provision of Title IV-B is the most recent source of funding for child welfare interventions. According to the U.S. Children’s Bureau:

> The primary goals of Promoting Safe and Stable Families (PSSF) are to prevent the unnecessary separation of children from their families, improve the quality of care and services to children and their families, and ensure permanency for children by reuniting them with their parents, by adoption or by another permanent living arrangement. States are to spend most of the funding for services that address: family support, family preservation, time-limited family reunification and adoption promotion and support.

A total of $381.3 million was allocated to the states in fiscal year 2016 in the form of block grants.

The second significant source of federal funding—and, by far, the most substantial—is Title IV-E of the Social Security Act. Title IV-E, created in 1980, is targeted exclusively for the costs of placing children into foster care, administering agencies that place and supervise children, and training the workforce that manages
Creating an Effective Child Welfare System

CHILD WELFARE

impacts. The Adoption and Safe
welfare systems. Even changes in
values and function of child
have yet to substantially influence
vasive press coverage of tragedies
lion-dollar settlements, and per-
tort actions resulting in multimil-
their clients. Class-action suits, civil
ready and willing to change than
Child welfare systems are no more
don't.

I often tell my social-work students
who want to work in the child wel-
standard child welfare
joke: “How many social workers
does it take to change a lightbulb?”
“One, if the lightbulb sincerely
wants to be changed.” I first raised
the point of this joke 20 years ago.21
Most of our child welfare interven-
tions would work only for those par-
ents and caregivers who are ready
for change. The stark reality is that
caregivers who maltreat their chil-
dren are no more willing to change
their behaviors than are smokers, or
those who are overweight, or those
of us who should use sunscreen but
don’t.

Child welfare systems are no more
ready and willing to change than
their clients. Class-action suits, civil
t tort actions resulting in multimil-
ion-dollar settlements, and per-
vasive press coverage of tragedies
have yet to substantially influence
the values and function of child
welfare systems. Even changes in
the laws have had mostly modest
impacts. The Adoption and Safe
Families Act of 199722 did seem to
result in an increase in adoptions
out of the foster-care system (from
37,000 in 1998 to 50,400 in 2014)
and a decrease in the average time
that children spend in foster care
(from 32.5 months in 1998 to 20.8
months in 2014).

The law also had an “aggravated cir-
cumstance” provision, which allows
states or counties to bypass reason-
able efforts to keep families together
and go directly to the termination of
parental rights if a court determines
that aggravated circumstances exist.
Examples of aggravated circum-
stances:
• Abandonment, torture, chronic
abuse, and sexual abuse. The
parent murdered another child of
the parent.
• The parent committed voluntary
manslaughter of another child of
the parent.
• The parent aided or abetted,
attempted, conspired, or solicited
to commit such a murder or vol-
tuntary manslaughter.
• The parent committed a felony
assault that resulted in serious
bodily injury to the child or another
child of the parent.
• The parental rights of the parent
to a sibling of the child were ter-
minated involuntarily.

The real problem dogging the U.S.
child welfare system is not insufficient
funds but insufficient flexibility in
how the existing funds may be used.
Because many foster family agen-
cies are dependent on the admin-
istrative costs provided under Title
IV-E, it creates a perverse incentive
that punishes foster-care agencies for
having unfilled foster-care beds. No
wonder child welfare administrators,
even as they complain about insuf-
ficient funds, resist changes in Title
IV-E funding.

A plausible change could create
greater flexibility: the Family First
Prevention Services Act failed to pass
the U.S. Senate in 2016 and was rein-
troduced in January as H.R. 253, the
Family First Prevention Services Act
of 2017.20 This bill would transform
the open-ended entitlement of Title
IV-E into a block grant and provide
more flexibility in funding for state
child welfare agencies. It would also
provide funds for evidence-based
interventions. Not surprisingly, there
is opposition to the bill from parent
advocates as well as institutions that
would lose funding under the new
funding system for Title IV-E.

3 IS THE CHILD WELFARE SYSTEM CAPABLE OF CHANGING?

Unfortunately, the aggravated
circumstances provision of the law
is rarely applied by child welfare
systems.23 There is certainly reason
to be pessimistic about the American
child welfare system. Still, change
is possible. The crucial problems
are agreeing on who the proper
client of the system should be, how
to improve decision making, and
eliminating the perverse incentive
of current foster care funding. The
essential solutions, in my judgment,
are to focus on the child as the client,
to make the child’s safety and well-
being the goal of the system, and
to abandon clinical judgment as the
basis for critical and life-and-death
decisions. An overdue revision of
Title IV-E of the Social Security Act
will free up billions of dollars for the
child welfare system. In the end, the
lightbulb still must sincerely want
to be changed: that will remain the
challenge for systems that cling to
the belief that parents are the most
important clients.
Endnotes


4 Ibid.

5 Ibid., chap. 3.


7 95 Stat. 500, Public Law 96-272.


11 Gelles, Out of Harm’s Way,” chap. 4.

12 Ibid., chap. 5.


14 For a more complete discussion, see Gelles, Out of Harm’s Way, chap. 5.


18 PL 74-271.


20 H.R. 253.


22 PL 105-89.

POLICING
“Predictive policing used to be the future,” said career cop William Bratton, “and now it is the present.” In mid-May 2015, Bratton—the visionary former chief of police in Boston, New York City, Los Angeles, and currently again top cop in New York—was talking about his early days as an officer in Boston, about what worked and what didn’t, and about what can work better in the future. That future will involve predictive policing, which Bratton is bringing to New York (a pilot program was launched last summer).

Predictive policing, which Bratton helped develop when he headed the Los Angeles Police Department during the 2000s, seeks not just to fight crime but to anticipate and prevent it. It uses cutting-edge technology and Big Data—some of which comes from past analysis and some of which is new, streaming in real time to an onboard computer in a patrol car—to identify high-risk areas, which precincts can then flood with police. The aim is not to make arrests but to deter crime before it occurs. Predictive policing relies crucially on community engagement—it can work only when the police are seen as part of the neighborhood, rather than as an occupying presence. At a time when police-community relations are frayed and many cities face rising violent-crime rates as well as renewed concern about terrorist threats, the approach may provide a better way forward.

Currently, about a dozen American cities—including Los Angeles as well as Santa Cruz, Atlanta, Georgia, and Tacoma, Washington—are using PredPol, a leading predictive-policing software and analytics program. Many of these cities rolled out the system over the past few years, and they are seeing positive, and sometimes dramatic, results. In Los Angeles, a nearly two-year study by UCLA crime scholars and law-enforcement officials, released this past fall, found that PredPol successfully predicted—and prevented—twice as much crime as human crime analysts did. The LAPD is now using PredPol in 14 of its 21 divisions.

“Every police department in cities of 100,000 people and up,” says criminologist Craig Uchida, “will be using some form of predictive policing in the next few years.” Like every innovation, the method has advocates and critics. Predictive policing strikes its detractors as potentially Orwellian law enforcement; but at its best, it aspires to something quite the opposite—a return, albeit a high-tech one, to the days of police on the beat who knew their constituents and worked with them to keep neighborhoods safe.
The use of information to respond to crime has always been part of the history of policing, an essential part to solving crime after the fact,” Bratton explains. “When Sir Robert Peel [home secretary in early-nineteenth-century England] created the British Metropolitan Police Force, he had nine principles of policing, which were focused on the prevention of crime.”

Before Peel created the Met in 1838, London was policed by the Bow Street Runners, six officers who constituted the city’s first professional police force and solved crimes as a civic service—not, as professional “thieftakers” previously did, for a fee. The force was founded in 1749 by Henry Fielding, author of Tom Jones. It disbanded in 1839, the year after Peel created the Met as a citizen police presence and an alternative to a military force. The Met’s authority depended on public approval. To Peol, this meant that the police—nicknamed “Bobbies,” for Peel’s first name—had to behave respectfully, succeeding not through compulsion but through the willing cooperation of citizens. Bobbies traditionally didn’t carry firearms; when force was necessary, it was to be minimal. And the police were not the judiciary. Bobbies did not judge guilt or innocence, did not punish or avenge.

“The police are the public,” Peel said, “and the public are the police.” Peel believed that the Met proved itself not by the number of criminals it caught but by the absence of crime.

In the United States, policing took shape along Peelian principles: cops walking the city beat deterred crime by their presence. They got to know the neighbors and the neighborhoods, which were more stable than they became after World War II. For years, many cops claimed that the most realistic police show on TV was Barney Miller because it showed the everyday neighborhood problems that most officers dealt with: homicides were rare, and few cops ever had cause to pull a gun. In New York City, you drew your weapon only in life-threatening situations, and, when you did, you shot to kill. The most common response to shooting a dangerous criminal was to vomit. But Barney Miller was an exercise in nostalgia. It was produced just as the system was changing.

Bratton began his policing career in 1970. “It was a time of great transition,” he says. “We were coming out of the 1960s, that time of social turbulence.” Police had become reactive. Instead of a benign presence, designed to prevent crime, they spent more and more time in their squad cars, responding to calls. At a time when people were attacking police, throwing bricks off roofs, using guns and Molotov cocktails, and fomenting revolution in the streets, patrol cars were seen as safer—and more efficient, as budgets became stretched. Officers could cover a greater area more quickly in a car than on foot. Random patrols and responses to calls—counterpunching—supplanted an assigned beat. It was “an ultimately damaging refocusing of the police mission and responsibility,” as Bratton puts it.

The introduction of the 911 calling system, starting in the late 1960s, reinforced the trend to reactive policing, as did a technological revolution in everything from ballistics to serology. The lab replaced the street. Suddenly, cops had a new toolbox of crime-solving techniques. “Policing was professionalized,” Bratton said. There was “better training, better education, and more focus on solving crime than preventing it.”

But the new methods generated new problems: police departments found themselves swamped with the surge in calls coming in because of 911. They couldn’t keep up with the demand for their services. Making matters worse: to save money, cities began cutting back police forces around the country. The Boston PD went from 2,800 officers to 2,000. New York City’s force shrunk from 31,698 in 1970 to 22,588 in 1980, about two-thirds the size of the force today. The police who remained faced many new frustrations, especially with the growing legal emphasis on criminals’ civil rights. The Supreme Court’s 1964 ruling in Escobido v. Illinois (establishing the right of suspects to have a lawyer present during police questioning) and the 1966 ruling in Miranda v. Arizona (requiring police to read suspects their rights before they can be questioned) transformed policing, as did modern legal interpretations of the Exclusionary Rule, which deems inadmissible in court any evidence that police collect in violation of a suspect’s constitutional rights. Cops were overwhelmed, too, by the growth of trafficking in hallucinogens and the various forms of cocaine and by the increasing numbers of young people drawn into the drug culture—abandoning homes and schools, hitch-hiking across the country, and living on the streets, in communes, and in squats. Haight-Ashbury and the East Village replaced the traditional skid row. Things worsened during the 1980s. “We were being briefed [in Boston] on this new type of cocaine, crack, that had decimated New York,” Bratton said, “and we were starting to see signs in Massachusetts within a short period of time: whoosh, it was like one of these tsunamis.” Crime numbers climbed to frightening new highs.

Deinstitutionalization compounded the problem. A flood of homeless mentally ill people ended up on the streets, degrading neighborhoods further. Every day, people were confronted with evidence of a crumbling social contract. By the end of the 1980s, Bratton observes, the reactive approach dominated police methods, which focused on improving performance in two areas: responding to 911 calls and investigating major crimes. Police neglected a third: preventing disorder. Urban quality of life was deteriorating, with many neighborhoods overrun with litter, crack vials, prostitution, street people, panhandlers, squeegee men—and broken windows.

In March 1982, George L. Kelling and James Q. Wilson’s soon-to-be-famous article “Broken Windows,” about policing that focused on quality-of-life crimes, appeared in The Atlantic. Kelling and Wilson’s theory was that tolerating disorder soon bred the conditions for more serious crimes, as lawbreakers felt emboldened. In the
1990s, during the early years of Rudy Giuliani’s mayoralty, when Bratton began his first tenure as New York’s police commissioner, a new “Broken Windows” philosophy of policing, with Bratton and Kelling as leading advocates (often in the pages of City Journal), replaced the more reactive approach. Broken Windows targeted misdemeanor offenses and prioritized neighborhood order as an important goal in itself. “Working closely with the community to identify what was creating fear,” Bratton says, police operated on “the recognition that fear was being created not only by serious crime but by quality-of-life crimes.” In arresting low-level offenders, moreover, police often found that they were wanted for more serious felonies. Analysis of “root causes” was left to sociologists. “We believed that the causes of crime were not racism, poverty, the economy, demographics,” says Bratton. “They could be significant influences if not dealt with, but the real causes of crime are people—criminals.”

The Broken Windows theory was reinforced by CompStat, the crime analysis and accountability system that Jack Maple created and instituted in the New York Police Department in 1994, when Bratton began his first tenure as the city’s police commissioner. CompStat, the forerunner of predictive policing, helped Bratton and the NYPD transform a city that had been overrun for decades by crime and disorder. “A wide variety of changes in how patrol resources were deployed in New York City were at the center of the regime initiated by William Bratton and crime analyst Jack Maple and continued for years afterward,” observes Franklin E. Zimring in The City That Became Safe. “Crime analyses and mapping were used to target particular problem areas where patrol, detective, and narcotics units might be emphasized.”

How best to use the large amounts of information the police were processing? “We began to use intelligence-led policing to focus on the prevention of crime,” Bratton said. “And, increasingly, as we got more capacity to get that information in a timely fashion [real time], we were able to stop patterns and trends after the second, third, or fourth incident rather than catching them when they’ve committed 20 or 30 or 40. We had found a way to prevent crime by stopping perpetrators after the second or third act.”

The result of the NYPD’s innovations is well known to City Journal readers: between 1993 and 2000, a drop of 57.26 percent took place in the seven major crime categories, from murder and rape to burglary and auto theft. New York’s success was so striking that many cities across the nation adopted the new methods. Not only was technology offering unprecedented amounts of data; it also was developing algorithms that could analyze the data more quickly, and get the analysis to the cop in the patrol car. Two decades after these practices first became systematized, the building blocks were in place for the next phase of the policing revolution: predictive policing.

The PredPol system works by looking at a specific geographical area. A precinct gets divided up into 500-by-500-foot “boxes,” each about the size of a city block. A box thought likely to be a locus of crime—because of its history, location, current activity, and other factors—becomes a “hot box.” A computer analyzes what crimes have occurred there in the past, what is happening in real time, and other inputs. Even the weather can be a factor: when it’s below freezing, for example, fewer potential criminals come out on the streets.

What sites in the area—schools, strip clubs, bars, malls—might serve as magnets for potential burglars? The cops look at an “intersection of risk factors,” said Megan Yerxa, a crime analyst at the Tacoma, Washington, police department, which fully deployed PredPol in October 2013. They have to look at “a terrain that shows hot areas,” such as places where two cultures rub up against each other, as in New York City, where the Upper East Side abuts Spanish Harlem. Once the hot spots are identified, the police visit them two or three times during a tour, making their presence felt and, by doing so, deterring crime. “PredPol doesn’t take into [account] socioeconomic factors,” Yerxa says. It avoids targeting a particular race or income group, and the statistics tend to hold true for all kinds of neighborhoods. For example, according to Yerxa, residential burglary frequently occurs near high schools and middle schools, and not just in bad neighborhoods. “If a home is broken into in an affluent area,” explains UCLA anthropology professor Jeff Brantingham, PredPol’s co-creator, people mistakenly think that it’s an “outsider,” someone from a poorer neighborhood preying on richer people. But “it’s most likely the teenager down the street trying to impress friends.” Because PredPol focuses not on a city, not on a borough, and not even on a neighborhood, but on a block or a few blocks, it is very specific and localized.

The evidence suggests that crime is predictable, at least statistically, because criminals follow patterns. They tend to go back to work in the same area, the same house or business, because they’re familiar with the setup—and they got away with it the first time. The nearby risk of one burglary, for instance, tends to be 600 feet from the first crime over a period of 14 to 28 days. “An offender hits a neighborhood three times,” Yerxa said—and then tends to go to the next neighborhood over. “They travel to avoid detection, but they don’t travel too far.” The question she asked herself was: “Do the last 28 days predict the next seven days?” How do they predict it? And is there a way to intervene to interrupt the repeating pattern of criminal behavior without violating someone’s civil rights?

Brantingham and his PredPol cofounder George Mohler of Santa Clara University thought it was possible to predict certain crimes, using a system similar to the one that forecasts earthquake aftershocks. Initial crimes are often followed by an “aftercrime”—for instance, a car theft can be followed by a robbery in which the stolen car is used. They developed an algorithm to predict future crime...
locations. Departments using PredPol look at what has happened recently in a hot box—say, a series of house break-ins before dawn. The local precinct may send patrol cars and beat cops into that box.

If the cops spot a known offender—one with a record of that particular crime at that time of day—and the KO is casing a house, that gives the officer probable cause to stop and engage the suspect and—possibly—to make an arrest. “I’ll stop and talk to a guy,” said Los Angeles Foothill Division assistant watch commander Sergeant Scott Kennedy, “and engage him.” Kennedy tries to create a consensual, not a confrontational, encounter. This is not a dragnet, scooping up anyone who looks suspicious. It is targeted—and it’s used only when there is clear probable cause. “Positive outreach,” according to Tacoma’s senior crime analyst Jackie Shelton, Yerxa’s colleague, “is as important as deterrence.” Community-based services, she says, are important if PredPol is to be effective.

PredPol is based on universal principles. “I got hooked on Moneyball and sabermetrics (Bill James’s revolutionary analytic tool for judging baseball players’ performance),” says Craig Uchida, the criminologist who is also president of Justice and Security Strategies—and a member of the Society of Baseball Researchers. When he worked for the Department of Justice, Uchida oversaw grants to develop predictive-policing methods. He noticed the similarity between baseball and police work. “Look at baseball,” he says. “Baseball collects lots of information about everything. You can create different algorithms for different baseball players and how they perform”—and you can do the same for crime.

Cops have always depended on intuition, Uchida says. Now they can “back up intuition with information—have cops talk with people on the streets to find out what is really happening. Write it down. Turn it into statistics, and then—using new technology that we now have—crunch the shared experience.” The outcome depends on good data. “This is a shift in how you do police work,” he says.

When Bratton called Uchida in 2008 to come out to Los Angeles and work with him and Sean Malinowski, the project manager for the LAPD’s predictive-policing initiative and now chief of staff to Bratton’s successor in L.A., Charlie Beck, Uchida says, “I had no idea what he was talking about. Like everyone else, I was skeptical. But there was something about the LAPD that I liked. There seemed to be something there.” He met Jeff Brantingham, whose parents, Uchida said, were “stars” in the criminology field, formulatores of the “environmental crime theory” that looks at what situations cause or prevent crime. It was like the gathering of the gang in The Magnificent Seven: Uchida, Brantingham, Malinowski, all working under Bratton—a happy coincidence of people and project that Bratton creates wherever he goes.

By July 2011, Mohler and Brantingham’s PredPol company was supplying increasingly sophisticated and effective algorithms to the LAPD and about a dozen other police departments. At first, though, the process was rudimentary. “George [Mohler] and Jeff [Brantingham] had a great concept,” notes Santa Cruz deputy police chief Steve Clark, “but they didn’t know how to make it relevant in the police world.” Clark and his analysts would gather crime statistics, enter them as Excel columns, and give each case a geographical coordinate. Brantingham and Mohler then cranked out a map with hot boxes. They refined the algorithms over time.

Clark didn’t understand the math, but he knew what it tried to capture: pockets of crime in three key areas—burglaries, theft from vehicles, and theft of vehicles. Those crimes, he said, “are the ones we felt we could [predict] with a good degree of accuracy. Instead of chasing yesterday’s crime, we used yesterday’s crime to predict where today’s crime is going to happen.”

The Santa Cruz PD started testing the predictive accuracy of PredPol. “In some cases, we were 60 percent predicting where or when,” Clark said. “It blew my mind.” Other jurisdictions were seeing increases in crime; Santa Cruz saw reductions. “In our first six months, we saw a drop of 28 percent in burglary compared with the same locations and same time in the previous year,” he observes. They started knocking out “the low-hanging fruit,” he says. “If you catch one burglar, he can’t knock over a bunch of houses” later. Future reductions flow from already-reduced numbers. And crime continued to drop. Burglaries fell by 7 percent, robberies by 27 percent. Arrests increased by 56 percent, while recovery of stolen cars increased by 22 percent.

Like Santa Cruz, other cities using PredPol have been getting promising results—an overall reduction in crime compared with other boxes in the city and in the same box a year ago. The results were immediate in Los Angeles. In the L.A. Foothill Division, in the four months following the start of the program in 2012, there was a drop of 13 percent, compared with an increase of 0.4 percent in the rest of the city. Between May 2012 and November 2012, during the LAPD pilot program, there was a 12 percent combined reduction in auto theft, burglary, and theft—the three crimes that L.A.’s PredPol tracks. Burglary alone saw a drop of 25 percent, compared with the same period in the same area a year earlier. In 2014, the Foothill Divi-
Big Data on the Beat—Predictive Policing Has Arrived

The police can make cities safe, but there is always a cost. Freedom or security: it’s a balance that we weigh every day. At what point does protecting society run the danger of infringing excessively on individual rights, or, more broadly, changing the nature of our democratic republic? And are there reasonable alternatives to these choices?

“Predictive policing has limited use,” argues Constance Rice, a civil rights activist, who cofounded and codirects the Advancement Project of Los Angeles. She finds the name “misleading. It’s not like they have a crystal ball, and they can predict where crime can be.” She’s skeptical that the time and resources devoted to predictive policing are worth it. She adds, “There can be civil liberties repercussions to anything the police do.”

“Predictive policing is not new,” says Richard Berk of the Wharton School at the University of Pennsylvania. “If you go back to the old TV shows—Dragnet—there’d be a map of the city, and there would be pins, and the pins would be concentrated in one area more than another. That’s been around since the 1950s.” According to Berk, predictive policing is just a fancier version of these efforts. It’s a quantitative, not a qualitative, difference—though in most fields, quantitative differences matter. “The additional feature is the mathematics, which helps interpret those pins on the map,” Berk admits, and he concedes that predictive policing may not just displace crime but reduce it. But he doubts “whether it’s as dramatic as some claim.”

There are “three potential problems with predictive policing,” according to Chris Calabrese of the American Civil Liberties Union. “First, garbage

During roll call one evening, before Officers Allen and Bayer (their names have been changed for this article) had left the Los Angeles Foothill Division station house, they’d been given the daily printout of hot boxes in their precinct. In the watch commander’s office, a wall monitor displayed all the 500-by-500-foot boxes covering the entire division, with the hot boxes highlighted in red. Nearby, on the wall, eight videos showed live CCTV views of Van Nuys and some of the other main streets in the precinct—cars passing, an occasional pedestrian.

The two officers were not responding to a particular call—they were on “additional” patrol, covering predictive box M at Laurel Canyon and Mercer—one of the hot boxes in the 47.3 square miles of the Foothill Division. A mixed racial and economic community of 250,000 people, with some light industry, it consists of well-maintained one-story houses with porches and neatly kept front yards. Thirty gangs, many from other neighborhoods, show up at night. But the fight is gone. The gang members act more like the commuters they are. Fifteen to 20 years ago, the area saw frequent gang fights and shootings, stolen and torched cars, particularly around Hubert Humphrey Park (which, today, holds well-lit nighttime soccer games). Now, the area is safer. Kids who in previous eras may have been loitering on the corners tend to be inside more these days, playing video games.

In their patrol car, Allen and Bayer monitored their laptop, which connected them to the station house and various databases, including those containing car license-plate numbers and criminal records. As they drove through the Pacoima streets, they scanned the area for anything suspicious. “If I see someone walking in shadows instead of under streetlights, I might be suspicious,” said Sergeant Scott Kennedy, Foothill’s assistant watch commander. “At night, most people want to stay in the light.”

What else can make a cop suspicious? Someone holding his or her hand waist-level—possibly ready to pull a gun or drop a bag of dope. On a corner, guys glancing over one another’s shoulders as they talk, as if on the lookout; a car hugging the curb—someone who seems to be trolling, not going anywhere; and, of course, cars that match reports of stolen vehicles. Cops can follow bulletins on their onboard computers while responding to traditional calls from the precinct house.

At the hot box near Laurel and Mercer, Allen and Bayer spotted a car with five occupants—three females and two males, two of whom the cops recognized as gang members. Because the gang members had troubled histories, the cops called for backup from two officers working other hot boxes. When Allen and Bayer stopped the car, they found that one of the two guys in the backseat was carrying a loaded, snub-nosed .38 caliber pistol. A double-barrel 12-gauge shotgun lay on the rear floorboard. The male passenger copped to both weapons, and one of the female passengers was arrested for possession of a controlled substance. Back at the station, the officers checked in with officers from the Foothill Gangs Unit, who, it turns out, were looking for the gun suspect. The stop may lead to additional weapons charges and arrests.

“This was a great arrest,” Malinowski says, “a terrific example of being in the right place at the right time and of the seamless collaboration between detectives, patrol, and specialized units that we’ve got going here. We know from our experience that five gang members in a car with a pistol and a shotgun were probably going to end up in a shooting at some point, right? I don’t think I’m exaggerating when I say that we saved ourselves a shooting and a couple of victims.”

The week ended with only 42 crimes in the whole division,” says Malinowski, who was captain of the Foothill Division at the time. “That’s in an area 50 square miles with a population of 250,000.” On Thursday of that week, February 13, the Foothill Division recorded no crime. Zero. “So,” Malinowski said, “something is working, right?” It is these striking results that the recent UCLA study captures and that led the LAPD to expand the program.
in, garbage out. If you put bad data in, you get bad results. By bad data, I mean historical inequities.” Calabrese focuses on stop-and-frisk, which he seems to equate with predictive policing, though most advocates of the latter are careful to distinguish between these approaches. The second problem, Calabrese said, “is transparency: how are these algorithms created, can you audit them, or are they a black box where the computer tells you where you focus, and there’s no accountability? Who is vetting the algorithm?” This concern may be more reasonable, since the brain of a computer is mysterious to most cops. The police must educate themselves about the algorithms used and demand clear explanations from those creating them. Perhaps citizen-review boards with appropriate expertise could help vet the algorithms. “Third,” says Calabrese, “what types of enforcement are being used?” Are the police looking at “where crimes were reported in the last few months, [or] are they looking on Facebook pages? Are they looking at friends of people who have been targeted on Facebook?” Would the ACLU be willing to work with PredPol? I ask. Calabrese hesitates: “Our work [is] to monitor civil rights.” But “our door is always open.”

“My fear is that these programs are creating an environment where the police can show up at anyone’s door at any time for any reason,” says Hanni Fakhoury, an Electronic Frontier Foundation lawyer. Fakhoury cites an incident in which the Chicago PD warned Robert McDaniel, 22, that he was on a “heat” list generated by an algorithm developed at the Illinois Institute of Technology to identify those most prone to violence and that he had no violent criminal record. “This is what civil rights activists are concerned about,” concedes the LAPD’s Malinowski. The worry is that police will create, as one headline suggested, a “real Minority Report computer to predict crimes,” referring to the dystopian film and its concept of “pre-crime,” in which individuals are arrested based on offenses that they supposedly would commit in the future. A number of predictive-policing software and analytics systems have been developed—IBM Modeler, Bair Analytics, and ESRI Mapping, among others—that do tend to focus on individuals. That could lead to constitutional or civil rights challenges in the future.

The PredPol police departments are different. None would generate a “heat” list such as the one used in Chicago. Again, PredPol focuses on geography—“place-based policing,” as Tacoma assistant police chief Peter Crimmins describes it. “The Constitution doesn’t go away,” adds Uchida. “We made a conscious decision that we weren’t going to target individuals. PredPol has stayed away from individuals and will continue to do that.” To prevent the new era of Big Data from becoming authoritarian—to balance security with freedom—Bratton, who is sensitive to the danger, says that “it’s also important to make sure to have good people elected and appointed to office.” He continues: “The trust on which democracy itself rests is: you have standards of oversight and the ability to attract the right people into these positions of great trust. Isn’t that the whole foundation of what democracy is based on? We give up certain freedoms to other individuals, who have the power—as in the case of the police—to control your behavior. The challenge for the police is that it be done constitutionally. Legally. Respectfully. Compassionately. That it be done consistently. By consistently, I mean not policing differently in a poor neighborhood than in a rich neighborhood.”

The best defense against misuse of the police is citizen vigilance. When we give up our duties to police our neighbors to a professional force, we are not surrendering our responsibility to monitor those who act on our behalf. It seems inevitable that some form of predictive policing is on the way. We can’t put the technology of policing back in the bottle any more than the Luddites could stop power looms. We’ll need to decide what forms the system should take—especially with the mounting danger to cities from terrorism. “As criminal and terrorist threats evolve,” Malinowski said a few days after the Paris and Mali terrorist attacks in November 2015, “law enforcement will need to continue to develop ways to use the vast amount of data available to analyze, measure, evaluate, and forecast risks in order to mitigate the impact of crime and terror events.” Brantingham concurs. “Terrorist events will remain difficult to predict, as they are rare in a statistical sense,” he says, “but the convergence between crime and terrorism indicates that geographically based predictive policing that prevents crime by disrupting opportunities can also interfere with the freedom of action necessary in the planning of terrorist events. PredPol makes no assumptions about what specific environmental and social conditions drive crime or terrorism—and therefore, it is less subject to the biases that such assumptions can entail.”

The most optimistic scenario is that the technological revolution—through systems like PredPol—re-creates the best of traditional beat-cop policing. Officers will get to know their community and will, as in the PredPol model, incorporate community outreach. Police departments will do more with tighter budgets and fewer resources while also protecting individual rights. “For every 10 percent reduction of officers,” said Tacoma’s Yerxa, “there is a 5 percent increase in crime.” PredPol allows cops to reinvent Robert Peel’s preventive policing—to return to the humane principles of the 1830s with twenty-first-century tools. Bratton recently brought Malinowski to New York to consult on building a predictive-policing capability for the NYPD—and he envisions more interaction among the three great police departments in New York, L.A., and London. “We’re testing both PredPol and a new system, Hunchlab.” Bratton explains. “We haven’t committed to either. In the meantime, we have developed our own predictive model that we are field-testing against the other two.”

Now 68, with nearly a half-century in
law enforcement behind him, Bratton remains as committed as ever. “It’s my heart, soul, and passion,” he says. “The importance of policing. No one can do it better than us, but we can’t do it alone.”

Night. The L.A. Foothill Division patrol car passed IncoTech Aerospace Fasteners, near what used to be gang turf, along Van Nuys Boulevard, past Rosie’s Tax and Plus. San Fernando Road—past Chabelita’s Restaurant, not far from where dealers used to peddle heroin. Best Buy, Costco, Lowe’s, a Subway, a Panda Express. Hillery T. Broadous Elementary School. One story ranch houses, beach chairs on porches, front doors flanked by lights. Walls that used to be tagged up that are now free of graffiti. A few stables. Not long ago, someone was arrested for sex with a horse.

On his patrol-car computer, Sergeant Scott Kennedy, the Foothill Division assistant watch commander, punched in a code, letting the precinct know that he was entering a hot box. Kennedy slowed as he skirted Ritchie Valens Park. Hubert Humphrey Memorial Park: under banks of lights, kids played a soccer game. Paxton Park. Near the intersection of two freeways—210 and 118—Kennedy stopped and got out.

When Kennedy first joined the department, kids would push cars into the dry channel and set them on fire. Gangs fought over turf. There were muggings, murders. Thirty years ago, in most cities, when a cop came around the block, people alerted one another by whistling—just as at the beginning of West Side Story. Now, at night, you may see a twinkling of light spreading outward as people hit their cell phones to alert friends.

Kennedy walked into a tunnel, an underpass, still covered with graffiti. As long as one gang didn’t spray over or cross out another gang’s tag, everything was cool. The tunnel smelled damp. An empty beer bottle. A used condom. At the end, Kennedy stopped, looked around. At the tunnel’s entrance, backlit, a shadow flicked past and then returned, joined by three other shadows and then two more. Six, seven bodies blocked the way back to his patrol car. Kennedy kept his eyes on the shadows at the end of the tunnel as he approached them. He stood up straighter. His right hand drifted down toward his holstered gun. But—except for that slight body adjustment—he made a point of being open, affable, as he greeted the people at the end of the tunnel.

“How’s it going? Nice night, huh? Everything okay?” Any tension was defused. The group dispersed. Kennedy returned to his patrol car.

Back at the Foothill station house, Malinowski asked me how the night went.


“Yeah,” Malinowski grinned. “That’s the story.”