pay less attention to WMDs like value-added models. Or better yet, jettison them entirely.

At around the same time, New York governor Andrew Cuomo's education task force called for a four-year moratorium on the use of exams to evaluate teachers. This change, while welcome, does not signal a clear rejection of the teacher evaluation WMDs, much less a recognition that they're unfair. The push, in fact, came from the parents, who complained that the testing regime was wearing out their kids and taking too much time in the school year. A boycott movement had kept 20 percent of third through eighth graders out of the tests in the spring of 2015, and it was growing. In bowing to the parents, the Cuomo administration delivered a blow to value-added modeling. After all, without a full complement of student tests, the state would lack the data to populate it.

Tim Clifford was cheered by this news but still wary. "The opt-out movement forced Cuomo's hand," he wrote in an e-mail. "He feared losing the support of wealthier voters in top school districts, who were the very people who most staunchly supported him. To get ahead of the issue, he's placed this moratorium on using test scores." Clifford fears that the tests will be back.

Maybe so. And, given that value-added modeling has become a proven tool against teachers' unions, I don't expect it to disappear anytime soon. It's well entrenched, with forty states and the District of Columbia using or developing one form of it or another. That's all the more reason to spread the word about these and other WMDs. Once people recognize them and understand their statistical flaws, they'll demand evaluations that are fairer for both students and teachers. However, if the goal of the testing is to find someone to blame, and to intimidate workers, then, as we've seen, a WMD that spews out meaningless scores gets an A-plus.

8

COLLATERAL DAMAGE

Landing Credit

Local bankers used to stand tall in a town. They controlled the money. If you wanted a new car or a mortgage, you'd put on your Sunday best and pay a visit. And as a member of your community, this banker would probably know the following details about your life. He'd know about your churchgoing habits, or lack of them. He'd know all the stories about your older brother's run-ins with the law. He'd know what your boss (and his golfing buddy) said about you as a worker. Naturally, he'd know your race and ethnic group, and he'd also glance at the numbers on your application form.

The first four factors often worked their way, consciously or not, into the banker's judgment. And there's a good chance he was more likely to trust people from his own circles. This was only human. But it meant that for millions of Americans the predigital status quo was just as awful as some of the WMDs I've been describing. Outsiders, including minorities and women, were routinely locked out. They had to put together an impressive financial portfolio—and then hunt for open-minded bankers.

It just wasn't fair. And then along came an algorithm, and things improved. A mathematician named Earl Isaac and his engineer friend, Bill Fair, devised a model they called FICO to evaluate the risk that an individual would default on a loan. This FICO score was fed by a formula that looked only at a borrower's finances—mostly her debt load and bill-paying record. The score was color blind. And it turned out to be great for the banking industry, because it predicted risk far more accurately while opening the door to millions of new customers. FICO scores, of course, are still around. They're used by the credit agencies, including Experian, Transunion, and Equifax, which each contribute different sources of information to the FICO model to come up with their own scores. These scores have lots of commendable and non-WMD attributes. First, they have a clear feedback loop. Credit companies can see which borrowers default on their loans, and they can match those numbers against their scores. If borrowers with high scores seem to be defaulting on loans more frequently than the model would predict, FICO and the credit agencies can tweak those models to make them more accurate. This is a sound use of statistics.

The credit scores are also relatively transparent. FICO's website, for example, offers simple instructions on how to improve your score. (Reduce debt, pay bills on time, and stop ordering new credit cards.) Equally important, the credit-scoring industry is regulated. If you have questions about your score, you have

the legal right to ask for your credit report, which includes all the information that goes into the score, including your record of mortgage and utility payments, your total debt, and the percentage of available credit you're using. Though the process can be slow to the point of torturous, if you find mistakes, you can have them fixed.

Since Fair and Isaac's pioneering days, the use of scoring has of course proliferated wildly. Today we're added up in every conceivable way as statisticians and mathematicians patch together a mishmash of data, from our zip codes and Internet surfing patterns to our recent purchases. Many of their pseudoscientific models attempt to predict our creditworthiness, giving each of us so-called e-scores. These numbers, which we rarely see, open doors for some of us, while slamming them in the face of others. Unlike the FICO scores they resemble, e-scores are arbitrary, unaccountable, unregulated, and often unfair—in short, they're WMDs.

A Virginia company called Neustar offers a prime example. Neustar provides customer targeting services for companies, including one that helps manage call center traffic. In a flash, this technology races through available data on callers and places them in a hierarchy. Those at the top are deemed to be more profitable prospects and are quickly funneled to a human operator. Those at the bottom either wait much longer or are dispatched into an outsourced overflow center, where they are handled largely by machines.

Credit card companies such as Capital One carry out similar rapid-fire calculations as soon as someone shows up on their website. They can often access data on web browsing and purchasing patterns, which provide loads of insights about the potential customer. Chances are, the person clicking for new Jaguars is richer than the one checking out a 2003 Taurus on Carfax.com. Most scoring systems also pick up the location of the visitor's computer. When this is matched with real estate data, they can draw inferences about wealth. A person using a computer on San Francisco's Balboa Terrace is a far better prospect than the one across the bay in East Oakland.

The existence of these e-scores shouldn't be surprising. We've seen models feeding on similar data when targeting us for predatory loans or weighing the odds that we might steal a car. For better or worse, they've guided us to school (or jail) and toward a job, and then they've optimized us inside the workplace. Now that it might be time to buy a house or car, it's only natural that financial models would mine the same trove of data to size us up.

But consider the nasty feedback loop that e-scores create. There's a very high chance that the e-scoring system will give the borrower from the rough section of East Oakland a low score. A lot of people default there. So the credit card offer popping up on her screen will be targeted to a riskier demographic. That means less available credit and higher interest rates for those who are already struggling.

Much of the predatory advertising we've been discussing, including the ads for payday loans and for-profit colleges, is generated through such e-scores. They're stand-ins for credit scores. But since companies are legally prohibited from using credit scores for marketing purposes, they make do with this sloppy substitute.

There's a certain logic to that prohibition. After all, our credit history includes highly personal data, and it makes sense that we should have control over who sees it. But the consequence is that companies end up diving into largely unregulated pools of data, such as clickstreams and geo-tags, in order to create a parallel data marketplace. In the process, they can largely avoid government

oversight. They then measure success by gains in efficiency, cash flow, and profits. With few exceptions, concepts like justice and transparency don't fit into their algorithms.

Let's compare that for a moment to the 1950s-era banker. Consciously or not, that banker was weighing various data points that had little or nothing to do with his would-be borrower's ability to shoulder a mortgage. He looked across his desk and saw his customer's race, and drew conclusions from that. Her father's criminal record may have counted against her, while her regular church attendance may have been seen favorably.

All of these data points were proxies. In his search for financial responsibility, the banker could have dispassionately studied the numbers (as some exemplary bankers no doubt did). But instead he drew correlations to race, religion, and family connections. In doing so, he avoided scrutinizing the borrower as an individual and instead placed him in a group of people—what statisticians today would call a "bucket." "People like you," he decided, could or could not be trusted.

Fair and Isaac's great advance was to ditch the proxies in favor of the relevant financial data, like past behavior with respect to paying bills. They focused their analysis on the individual in question—and not on other people with similar attributes. E-scores, by contrast, march us back in time. They analyze the individual through a veritable blizzard of proxies. In a few milliseconds, they carry out thousands of "people like you" calculations. And if enough of these "similar" people turn out to be deadbeats or, worse, criminals, that individual will be treated accordingly.

From time to time, people ask me how to teach ethics to a class of data scientists. I usually begin with a discussion of how to build an e-score model and ask them whether it makes sense to use "race" as an input in the model. They inevitably respond that such a question would be unfair and probably illegal. The next question is whether to use "zip code." This seems fair enough, at first. But it doesn't take long for the students to see that they are codifying past injustices into their model. When they include an attribute such as "zip code," they are expressing the opinion that the history of human behavior in that patch of real estate should determine, at least in part, what kind of loan a person who lives there should get.

In other words, the modelers for e-scores have to make do with trying to answer the question "How have people like you behaved in the past?" when ideally they would ask, "How have *you* behaved in the past?"

The difference between these two questions is vast. Imagine if a highly motivated and responsible person with modest immigrant beginnings is trying to start a business and needs to rely on such a system for early investment. Who would take a chance on such a person? Probably not a model trained on such demographic and behavioral data.

I should note that in the statistical universe proxies inhabit, they often work. More times than not, birds of a feather *do* fly together. Rich people buy cruises and BMWs. All too often, poor people need a payday loan. And since these statistical models appear to work much of the time, efficiency rises and profits surge. Investors double down on scientific systems that can place thousands of people into what appear to be the correct buckets. It's the triumph of Big Data.

And what about the person who is misunderstood and placed in the wrong bucket? That happens. And there's no feedback to set the system straight. A statistics-crunching engine has no way to learn that it dispatched a valuable potential customer to call center hell. Worse, losers in the unregulated e-score universe have little recourse to complain, much less correct the system's error. In the realm of WMDs, they're collateral damage. And since the

whole murky system grinds away in distant server farms, they rarely find out about it. Most of them probably conclude, with reason, that life is simply unfair.

. . .

In the world I've described so far, e-scores nourished by millions of proxies exist in the shadows, while our credit reports, packed with pertinent and relevant data, operate under rule of law. But sadly, it's not quite that simple. All too often, credit reports serve as proxies, too.

It should come as little surprise that many institutions in our society, from big companies to the government, are on the hunt for people who are trustworthy and reliable. In the chapter on getting a job, we saw them sorting through résumés and red-lighting candidates whose psychological tests pointed to undesirable personal attributes. Another all-too-common approach is to consider the applicant's credit score. If people pay their bills on time and avoid debt, employers ask, wouldn't that signal trustworthiness and dependability? It's not *exactly* the same thing, they know. But wouldn't there be a significant overlap?

That's how the credit reports have expanded far beyond their original turf. Creditworthiness has become an all-too-easy stand-in for other virtues. Conversely, bad credit has grown to signal a host of sins and shortcomings that have nothing to do with paying bills. As we'll see, all sorts of companies turn credit reports into their own versions of credit scores and use them as proxies. This practice is both toxic and ubiquitous.

For certain applications, such a proxy might appear harmless. Some online dating services, for example, match people on the basis of credit scores. One of them, CreditScoreDating, proclaims that "good credit scores are sexy." We can debate the wisdom of linking financial behavior to love. But at least the customers of

CreditScoreDating know what they're getting into and why. It's up to them.

But if you're looking for a job, there's an excellent chance that a missed credit card payment or late fees on student loans could be working against you. According to a survey by the Society for Human Resource Management, nearly half of America's employers screen potential hires by looking at their credit reports. Some of them check the credit status of current employees as well, especially when they're up for a promotion.

Before companies carry out these checks, they must first ask for permission. But that's usually little more than a formality; at many companies, those refusing to surrender their credit data won't even be considered for jobs. And if their credit record is poor, there's a good chance they'll be passed over. A 2012 survey on credit card debt in low- and middle-income families made this point all too clear. One in ten participants reported hearing from employers that blemished credit histories had sunk their chances, and it's anybody's guess how many were disqualified by their credit reports but left in the dark. While the law stipulates that employers must alert job seekers when credit issues disqualify them, it's hardly a stretch to believe that some of them simply tell candidates that they weren't a good fit or that others were more qualified.

The practice of using credit scores in hirings and promotions creates a dangerous poverty cycle. After all, if you can't get a job because of your credit record, that record will likely get worse, making it even harder to land work. It's not unlike the problem young people face when they look for their first job—and are disqualified for lack of experience. Or the plight of the longtime unemployed, who find that few will hire them because they've been without a job for too long. It's a spiraling and defeating feedback loop for the unlucky people caught up in it.

Employers, naturally, have little sympathy for this argument.

Good credit, they argue, is an attribute of a responsible person, the kind they want to hire. But framing debt as a moral issue is a mistake. Plenty of hardworking and trustworthy people lose jobs every day as companies fail, cut costs, or move jobs offshore. These numbers climb during recessions. And many of the newly unemployed find themselves without health insurance. At that point, all it takes is an accident or an illness for them to miss a payment on a loan. Even with the Affordable Care Act, which reduced the ranks of the uninsured, medical expenses remain the single biggest cause of bankruptcies in America.

People with savings, of course, can keep their credit intact during tough times. Those living from paycheck to paycheck are far more vulnerable. Consequently, a sterling credit rating is not just a proxy for responsibility and smart decisions. It is also a proxy for wealth. And wealth is highly correlated with race.

Consider this. As of 2015, white households held on average roughly ten times as much money and property as black and Hispanic households. And while only 15 percent of whites had zero or negative net worth, more than a third of blacks and Hispanic households found themselves with no cushion. This wealth gap increases with age. By their sixties, whites are eleven times richer than African Americans. Given these numbers, it is not hard to argue that the poverty trap created by employer credit checks affects society unequally and along racial lines. As I write this, ten states have passed legislation to outlaw the use of credit scores in hiring. In banning them, the New York City government declared that using credit checks "disproportionately affects low-income applicants and applicants of color." Still, the practice remains legal in forty states.

This is not to say that personnel departments across America are intentionally building a poverty trap, much less a racist one. They no doubt believe that credit reports hold relevant facts that help them make important decisions. After all, "The more data, the better" is the guiding principle of the Information Age. Yet in the name of fairness, some of this data should remain uncrunched.

...

Imagine for a moment that you're a recent graduate of Stanford University's law school and are interviewing for a job at a prestigious law firm in San Francisco. The senior partner looks at his computer-generated file and breaks into a laugh. "It says here that you've been arrested for running a meth lab in Rhode Island!" He shakes his head. Yours is a common name, and computers sure make silly mistakes. The interview proceeds.

At the high end of the economy, human beings tend to make the important decisions, while relying on computers as useful tools. But in the mainstream and, especially, in the lower echelons of the economy, much of the work, as we've seen, is automated. When mistakes appear in a dossier—and they often do—even the best-designed algorithms will make the wrong decision. As data hounds have long said: garbage in, garbage out.

A person at the receiving end of this automated process can suffer the consequences for years. Computer-generated terrorism no-fly lists, for example, are famously rife with errors. An innocent person whose name resembles that of a suspected terrorist faces a hellish ordeal every time he has to get on a plane. (Wealthy travelers, by contrast, are often able to pay to acquire "trusted traveler" status, which permits them to waltz through security. In effect, they're spending money to shield themselves from a WMD.)

Mistakes like this pop up everywhere. The Federal Trade Commission reported in 2013 that 5 percent of consumers—or an estimated ten million people—had an error on one of their credit reports serious enough to result in higher borrowing costs. That's troublesome, but at least credit reports exist in the regulated side

of the data economy. Consumers can (and should) request to see them once a year and amend potentially costly errors.*

Still, the unregulated side of the data economy is even more hazardous. Scores of companies, from giants like Acxiom Corp. to a host of fly-by-night operations, buy information from retailers, advertisers, smartphone app makers, and companies that run sweepstakes or operate social networks in order to assemble a cornucopia of facts on every consumer in the country. They might note, for example, whether a consumer has diabetes, lives in a house with a smoker, drives an SUV, or owns a pair of collies (who may live on in the dossier long after their earthly departure). These companies also scrape all kinds of publicly available government data, including voting and arrest records and housing sales. All of this goes into a consumer profile, which they sell.

Some data brokers, no doubt, are more dependable than others. But any operation that attempts to profile hundreds of millions of people from thousands of different sources is going to get a lot of the facts wrong. Take the case of a Philadelphian named Helen Stokes. She wanted to move into a local senior living center but kept getting rejected because of arrests on her background record. It was true that she had been arrested twice during altercations with her former husband. But she had not been convicted and had managed to have the records expunged from government databases. Yet the arrest records remained in files assembled by a company called RealPage, Inc., which provides background checks on tenants.

For RealPage and other companies like it, creating and selling reports brings in revenue. People like Helen Stokes are not

* Even so, I should add, fixing them can be a nightmare. A Mississippi resident named Patricia Armour tried for two years to get Experian to expunge from her file a $40,000 debt she no longer owed. It took a call to Mississippi's attorney general, she told the *New York Times*, before Experian corrected her record.

customers. They're the product. Responding to their complaints takes time and costs money. After all, while Stokes might say that the arrests have been expunged, verifying that fact eats up time and money. An expensive human being might have to spend a few minutes on the Internet or even—heaven forbid—make a phone call or two. Little surprise, then, that Stokes didn't get her record cleared until she sued. And even after RealPage responded, how many other data brokers might still be selling files with the same poisonous misinformation? It's anybody's guess.

Some data brokers do offer consumers access to their data. But these reports are heavily curated. They include the facts but not always the conclusions data brokers' algorithms have drawn from them. Someone who takes the trouble to see her file at one of the many brokerages, for example, might see the home mortgage, a Verizon bill, and a $459 repair on the garage door. But she won't see that she's in a bucket of people designated as "Rural and Barely Making It," or perhaps "Retiring on Empty." Fortunately for the data brokers, few of us get a chance to see these details. If we did, and the FTC is pushing for more accountability, the brokers would likely find themselves besieged by consumer complaints—millions of them. It could very well disrupt their business model. For now, consumers learn about their faulty files only when word slips out, often by chance.

An Arkansas resident named Catherine Taylor, for example, missed out on a job at the local Red Cross several years ago. Those things happen. But Taylor's rejection letter arrived with a valuable nugget of information. Her background report included a criminal charge for the intent to manufacture and sell methamphetamines. This wasn't the kind of candidate the Red Cross was looking to hire.

Taylor looked into it and discovered that the criminal charges belonged to another Catherine Taylor, who happened to be born on the same day. She later found that at least ten other companies were tarring her with inaccurate reports—one of them connected to her application for federal housing assistance, which had been denied. Was the housing rejection due to a mistaken identity?

In an automatic process, it no doubt could have been. But a human being intervened. When applying for federal housing assistance, Taylor and her husband met with an employee of the housing authority to complete a background check. This employee, Wanda Taylor—no relation—was using information provided by Tenant Tracker, the data broker. It was riddled with errors and blended identities. It linked Taylor, for example, with the possible alias of Chantel Taylor, a convicted felon who happened to be born on the same day. It also connected her to the other Catherine Taylor she had heard about, who had been convicted in Illinois of theft, forgery, and possession of a controlled substance.

The dossier, in short, was a toxic mess. But Wanda Taylor had experience with such things. She began to dig through it. She promptly drew a line through the possible alias, Chantel, which seemed improbable to her. She read in the file that the Illinois thief had a tattoo on her ankle with the name Troy. After checking Catherine Taylor's ankle, she drew a line through that felon's name as well. By the end of the meeting, one conscientious human being had cleared up the confusion generated by web-crawling data-gathering programs. The housing authority knew which Catherine Taylor it was dealing with.

The question we're left with is this: How many Wanda Taylors are out there clearing up false identities and other errors in our data? The answer: not nearly enough. Humans in the data economy are outliers and throwbacks. The systems are built to run automatically as much as possible. That's the efficient way; that's where the profits are. Errors are inevitable, as in any statistical program, but the quickest way to reduce them is to fine-tune the algorithms running the machines. Humans on the ground only gum up the works.

This trend toward automation is leaping ahead as computers make sense of more and more of our written language, in some cases processing thousands of written documents in a second. But they still misunderstand all sorts of things. IBM's *Jeopardy!*-playing supercomputer Watson, for all its brilliance, was flummoxed by language or context about 10 percent of the time. It was heard saying that a butterfly's diet was "Kosher," and it once confused Oliver Twist, the Charles Dickens character, with the 1980s techno-pop band the Pet Shop Boys.

Such errors are sure to pile up in our consumer profiles, confusing and misdirecting the algorithms that manage more and more of our lives. These errors, which result from automated data collection, poison predictive models, fueling WMDs. And this collection will only grow. Computers are already busy expanding beyond the written word. They're harvesting spoken language and images and using them to capture more information about everything in the universe—including us. These new technologies will mine new troves for our profiles, while expanding the risk for errors.

Recently, Google processed images of a trio of happy young African Americans and its automatic photo-tagging service labeled them as gorillas. The company apologized profusely, but in systems like Google's, errors are inevitable. It was most likely faulty machine learning (and probably not a racist running loose in the Googleplex) that led the computer to confuse *Homo sapiens* with our close cousin, the gorilla. The software itself had flipped through billions of images of primates and had made its own distinctions. It focused on everything from shades of color to the distance between eyes and the shape of the ear. Apparently, though, it wasn't thoroughly tested before being released.

Such mistakes are learning opportunities—as long as the system receives feedback on the error. In this case, it did. But injustice persists. When automatic systems sift through our data to size us up for an e-score, they naturally project the past into the future. As we saw in recidivism sentencing models and predatory loan algorithms, the poor are expected to remain poor forever and are treated accordingly—denied opportunities, jailed more often, and gouged for services and loans. It's inexorable, often hidden and beyond appeal, and unfair.

Yet we can't count on automatic systems to address the issue. For all of their startling power, machines cannot yet make adjustments for fairness, at least not by themselves. Sifting through data and judging what is fair is utterly foreign to them and enormously complicated. Only human beings can impose that constraint.

There's a paradox here. If we return one last time to that '50s-era banker, we see that his mind was occupied with human distortions—desires, prejudice, distrust of outsiders. To carry out the job more fairly and efficiently, he and the rest of his industry handed the work over to an algorithm.

Sixty years later, the world is dominated by automatic systems chomping away on our error-ridden dossiers. They urgently require the context, common sense, and fairness that only humans can provide. However, if we leave this issue to the marketplace, which prizes efficiency, growth, and cash flow (while tolerating a certain degree of errors), meddling humans will be instructed to stand clear of the machinery.

• • •

This will be a challenge, because even as the problems with our old credit models become apparent, powerful newcomers are storming in. Facebook, for example, has patented a new type of credit rating, one based on our social networks. The goal, on its face, is reasonable. Consider a college graduate who goes on a religious mission for five years, helping to bring potable water to impoverished villages in Africa. He comes home with no credit rating and has trouble getting a loan. But his classmates on Facebook are investment bankers, PhDs, and software designers. Birds-of-a-feather analysis would indicate that he's a good bet. But that same analysis likely works against a hardworking housecleaner in East St. Louis, who might have numerous unemployed friends and a few in jail.

Meanwhile, the formal banking industry is frantically raking through personal data in its attempts to boost business. But licensed banks are subject to federal regulation and disclosure requirements, which means that customer profiling carries reputational and legal risk. American Express learned this the hard way in 2009, just as the Great Recession was gearing up. No doubt looking to reduce risk on its own balance sheet, Amex cut the spending limits of some customers. Unlike the informal players in the e-score economy, though, the credit card giant had to send them a letter explaining why.

This is when Amex delivered a low blow. Cardholders who shopped at certain establishments, the company wrote, were more likely to fall behind on payments. It was a matter of statistics, plain and simple, a clear correlation between shopping patterns and default rates. It was up to the unhappy Amex customers to guess which establishment had poisoned their credit. Was it the weekly shop at Walmart or perhaps the brake job at Grease Monkey that placed them in the bucket of potential deadbeats?

Whatever the cause, it left them careening into a nasty recession with less credit. Worse, the lowered spending limit would appear within days on their credit reports. In fact, it was probably there even before the letters arrived. This would lower their scores and drive up their borrowing costs. Many of these cardholders, it's safe to say, frequented "stores associated with poor repayments" because they weren't swimming in money. And wouldn't you know it? An algorithm took notice and made them poorer.

Cardholders' anger attracted the attention of the mainstream press, including the *New York Times*, and Amex promptly announced that it would not correlate stores to risk. (Amex later insisted that it had chosen the wrong words in its message and that it had scrutinized only broader consumer patterns, not specific merchants.)

It was a headache and an embarrassment for American Express. If they had indeed found a strong correlation between shopping at a certain store and credit risk, they certainly couldn't use it now. Compared to most of the Internet economy, they're boxed in, regulated, in a certain sense handicapped. (Not that they should complain. Over the decades, lobbyists for the incumbents have crafted many of the regulations with an eye to defending the entrenched powers—and keeping pesky upstarts locked out.)

So is it any surprise that newcomers to the finance industry would choose the freer and unregulated route? Innovation, after all, hinges on the freedom to experiment. And with petabytes of behavioral data at their fingertips and virtually no oversight, opportunities for the creation of new business models are vast.

Multiple companies, for example, are working to replace payday lenders. These banks of last resort cater to the working poor, tiding them over from one paycheck to the next and charging exorbitant interest rates. After twenty-two weeks, a $500 loan could cost $1,500. So if an efficient newcomer could find new ways to rate risk, then pluck creditworthy candidates from this desperate pool of people, it could charge them slightly lower interest and still make a mountain of money.

That was Douglas Merrill's idea. A former chief operating officer at Google, Merrill believed that he could use Big Data to calculate risk and offer payday loans at a discount. In 2009, he founded a start-up called ZestFinance. On the company web page, Merrill proclaims that "all data is credit data." In other words, anything goes.

ZestFinance buys data that shows whether applicants have kept up with their cell phone bills, along with plenty of other publicly available or purchased data. As Merrill promised, the company's rates are lower than those charged by most payday lenders. A typical $500 loan at ZestFinance costs $900 after twenty-two weeks—60 percent lower than the industry standard.

It's an improvement, but is it fair? The company's algorithms process up to ten thousand data points per applicant, including unusual observations, such as whether applicants use proper spelling and capitalization on their application form, how long it takes them to read it, and whether they bother to look at the terms and conditions. "Rule followers," the company argues, are better credit risks.

That may be true. But punctuation and spelling mistakes also point to low education, which is highly correlated with class and race. So when poor people and immigrants qualify for a loan, their substandard language skills might drive up their fees. If they then have trouble paying those fees, this might validate that they were a high risk to begin with and might further lower their credit scores. It's a vicious feedback loop, and paying bills on time plays only a bit part.

When new ventures are built on WMDs, troubles are bound to follow, even when the players have the best intentions. Take the case of the "peer-to-peer" lending industry. It started out in the last decade with the vision of borrowers and lenders finding each other on matchmaking platforms. This would represent the democratization of banking. More people would get loans, and at the same time millions of everyday people would become small-time bankers and make a nice return. Both sides would bypass the big greedy banks.

One of the first peer-to-peer exchanges, Lending Club, launched as an application on Facebook in 2006 and received funding a year later to become a new type of bank. To calculate the borrower's risk, Lending Club blended the traditional credit report with data gathered from around the web. Their algorithm, in a word, generated e-scores, which they claimed were more accurate than credit scores.

Lending Club and its chief rival, Prosper, are still tiny. They've generated less than $10 billion in loans, which is but a speck in the $3 trillion consumer lending market. Yet they're attracting loads of attention. Executives from Citigroup and Morgan Stanley serve as directors of peer-to-peer players, and Wells Fargo's investment fund is the largest investor in Lending Club. Lending Club's stock offering in December of 2014 was the biggest tech IPO of the year. It raised $870 million and reached a valuation of $9 billion, making it the fifteenth most valuable bank in America.

The fuss has little to do with democratizing capital or cutting out the middleman. According to a report in *Forbes*, institutional money now accounts for more than 80 percent of all the activity on peer-to-peer platforms. For big banks, the new platforms provide a convenient alternative to the tightly regulated banking economy. Working through peer-to-peer systems, a lender can analyze nearly any data it chooses and develop its own e-scores. It can develop risk correlations for neighborhoods, zip codes, and the stores customers shop at—all without having to send them embarrassing letters explaining why.

And what does that mean for us? With the relentless growth of e-scores, we're batched and bucketed according to secret formulas, some of them fed by portfolios loaded with errors. We're viewed not as individuals but as members of tribes, and we're stuck with that designation. As e-scores pollute the sphere of finance, opportunities dim for the have-nots. In fact, compared to the slew of WMDs running amok, the prejudiced loan officer of yesteryear doesn't look all that bad. At the very least, a borrower could attempt to read his eyes and appeal to his humanity.

9

NO SAFE ZONE

Getting Insurance

Late in the nineteenth century, a renowned statistician named Frederick Hoffman created a potent WMD. It's very likely that Hoffman, a German who worked for the Prudential Life Insurance Company, meant no harm. Later in his life, his research contributed mightily to public health. He did valuable work on malaria and was among the first to associate cancer with tobacco. Yet on a spring day in 1896, Hoffman published a 330-page report that set back the cause of racial equality in the United States and reinforced the status of millions as second-class citizens. His report used exhaustive statistics to make the case that the lives of black Americans were so precarious that the entire race was uninsurable.