1

Bomb Parts

What Is a Model?

It was a hot August afternoon in 1946. Lou Boudreau, the player-manager of the Cleveland Indians, was having a miserable day. In the first game of a doubleheader, Ted Williams had almost single-handedly annihilated his team. Williams, perhaps the game's greatest hitter at the time, had smashed three home runs and driven home eight. The Indians ended up losing 11 to 10.

Boudreau had to take action. So when Williams came up for the first time in the second game, players on the Indians' side started moving around. Boudreau, the shortstop, jogged over to where the second baseman would usually stand, and the second baseman backed into short right field. The third baseman moved to his left, into the shortstop's hole. It was clear that Boudreau, perhaps out of desperation, was shifting the entire orientation of his defense in an attempt to turn Ted Williams's hits into outs.

In other words, he was thinking like a data scientist. He had analyzed crude data, most of it observational: Ted Williams *usually* hit the ball to right field. Then he adjusted. And it worked. Fielders caught more of Williams's blistering line drives than before (though they could do nothing about the home runs sailing over their heads).

If you go to a major league baseball game today, you'll see that defenses now treat nearly every player like Ted Williams. While Boudreau merely observed where Williams usually hit the ball, managers now know precisely where every player has hit every ball over the last week, over the last month, throughout his career, against left-handers, when he has two strikes, and so on. Using this historical data, they analyze their current situation and calculate the positioning that is associated with the highest probability of success. And that sometimes involves moving players far across the field.

Shifting defenses is only one piece of a much larger question: What steps can baseball teams take to maximize the probability that they'll win? In their hunt for answers, baseball statisticians have scrutinized every variable they can quantify and attached it to a value. How much more is a double worth than a single? When, if ever, is it worth it to bunt a runner from first to second base?

The answers to all of these questions are blended and combined into mathematical models of their sport. These are parallel universes of the baseball world, each a complex tapestry of probabilities. They include every measurable relationship among every one of the sport's components, from walks to home runs to the players themselves. The purpose of the model is to run different scenarios at every juncture, looking for the optimal combinations. If the Yankees bring in a right-handed pitcher to face Angels slugger Mike Trout, as compared to leaving in the current pitcher, how much more likely are they to get him out? And how will that affect their overall odds of winning?

Baseball is an ideal home for predictive mathematical modeling. As Michael Lewis wrote in his 2003 bestseller, *Moneyball,* the sport has attracted data nerds throughout its history. In decades past, fans would pore over the stats on the back of baseball cards, analyzing Carl Yastrzemski's home run patterns or comparing Roger Clemens's and Dwight Gooden's strikeout totals. But starting in the 1980s, serious statisticians started to investigate what these figures, along with an avalanche of new ones, really meant: how they translated into wins, and how executives could maximize success with a minimum of dollars.

"Moneyball" is now shorthand for any statistical approach in domains long ruled by the gut. But baseball represents a healthy case study—and it serves as a useful contrast to the toxic models, or WMDs, that are popping up in so many areas of our lives. Baseball models are fair, in part, because they're transparent. Everyone has access to the stats and can understand more or less how they're interpreted. Yes, one team's model might give more value to home run hitters, while another might discount them a bit, because sluggers tend to strike out a lot. But in either case, the numbers of home runs and strikeouts are there for everyone to see.

Baseball also has statistical rigor. Its gurus have an immense data set at hand, almost all of it directly related to the performance of players in the game. Moreover, their data is highly relevant to the outcomes they are trying to predict. This may sound obvious, but as we'll see throughout this book, the folks building WMDs routinely lack data for the behaviors they're most interested in. So they substitute stand-in data, or proxies. They draw statistical correlations between a person's zip code or language patterns and her potential to pay back a loan or handle a job. These correlations are discriminatory, and some of them are illegal. Baseball models, for the most part, don't use proxies because they use pertinent inputs like balls, strikes, and hits.

Most crucially, that data is constantly pouring in, with new statistics from an average of twelve or thirteen games arriving daily from April to October. Statisticians can compare the results of these games to the predictions of their models, and they can see where they were wrong. Maybe they predicted that a left-handed reliever would give up lots of hits to right-handed batters—and yet he mowed them down. If so, the stats team has to tweak their model and also carry out research on why they got it wrong. Did the pitcher's new screwball affect his statistics? Does he pitch better at night? Whatever they learn, they can feed back into the model, refining it. That's how trustworthy models operate. They maintain a constant back-and-forth with whatever in the world they're trying to understand or predict. Conditions change, and so must the model.

Now, you may look at the baseball model, with its thousands of changing variables, and wonder how we could even be comparing it to the model used to evaluate teachers in Washington, D.C., schools. In one of them, an entire sport is modeled in fastidious detail and updated continuously. The other, while cloaked in mystery, appears to lean heavily on a handful of test results from one year to the next. Is that really a model?

The answer is yes. A model, after all, is nothing more than an abstract representation of some process, be it a baseball game, an oil company's supply chain, a foreign government's actions, or a movie theater's attendance. Whether it's running in a computer program or in our head, the model takes what we know and uses it to predict responses in various situations. All of us carry thousands of models in our heads. They tell us what to expect, and they guide our decisions.

Here's an informal model I use every day. As a mother of three, I cook the meals at home—my husband, bless his heart, cannot remember to put salt in pasta water. Each night when I begin to cook a family meal, I internally and intuitively model everyone's appetite. I know that one of my sons loves chicken (but hates hamburgers), while another will eat only the pasta (with extra grated parmesan cheese). But I also have to take into account that people's appetites vary from day to day, so a change can catch my model by surprise. There's some unavoidable uncertainty involved.

The input to my internal cooking model is the information I have about my family, the ingredients I have on hand or I know are available, and my own energy, time, and ambition. The output is how and what I decide to cook. I evaluate the success of a meal by how satisfied my family seems at the end of it, how much they've eaten, and how healthy the food was. Seeing how well it is received and how much of it is enjoyed allows me to update my model for the next time I cook. The updates and adjustments make it what statisticians call a "dynamic model."

Over the years I've gotten pretty good at making meals for my family, I'm proud to say. But what if my husband and I go away for a week, and I want to explain my system to my mom so she can fill in for me? Or what if my friend who has kids wants to know my methods? That's when I'd start to formalize my model, making it much more systematic and, in some sense, mathematical. And if I were feeling ambitious, I might put it into a computer program.

Ideally, the program would include all of the available food options, their nutritional value and cost, and a complete database of my family's tastes: each individual's preferences and aversions. It would be hard, though, to sit down and summon all that information off the top of my head. I've got loads of memories of people grabbing seconds of asparagus or avoiding the string beans. But they're all mixed up and hard to formalize in a comprehensive list.

The better solution would be to train the model over time, entering data every day on what I'd bought and cooked and noting the responses of each family member. I would also include parameters, or constraints. I might limit the fruits and vegetables to what's in season and dole out a certain amount of Pop-Tarts, but only enough to forestall an open rebellion. I also would add a number of rules. This one likes meat, this one likes bread and pasta, this one drinks lots of milk and insists on spreading Nutella on everything in sight.

If I made this work a major priority, over many months I might come up with a very good model. I would have turned the food management I keep in my head, my informal internal model, into a formal external one. In creating my model, I'd be extending my power and influence in the world. I'd be building an automated me that others can implement, even when I'm not around.

There would always be mistakes, however, because models are, by their very nature, simplifications. No model can include all of the real world's complexity or the nuance of human communication. Inevitably, some important information gets left out. I might have neglected to inform my model that junk-food rules are relaxed on birthdays, or that raw carrots are more popular than the cooked variety.

To create a model, then, we make choices about what's important enough to include, simplifying the world into a toy version that can be easily understood and from which we can infer important facts and actions. We expect it to handle only one job and accept that it will occasionally act like a clueless machine, one with enormous blind spots.

Sometimes these blind spots don't matter. When we ask Google Maps for directions, it models the world as a series of roads, tunnels, and bridges. It ignores the buildings, because they aren't relevant to the task. When avionics software guides an airplane, it models the wind, the speed of the plane, and the landing strip below, but not the streets, tunnels, buildings, and people.

A model's blind spots reflect the judgments and priorities of its creators. While the choices in Google Maps and avionics software appear cut and dried, others are far more problematic. The value-added model in Washington, D.C., schools, to return to that example, evaluates teachers largely on the basis of students' test scores, while ignoring how much the teachers engage the students, work on specific skills, deal with classroom management, or help students with personal and family problems. It's overly simple, sacrificing accuracy and insight for efficiency. Yet from the administrators' perspective it provides an effective tool to ferret out hundreds of apparently underperforming teachers, even at the risk of misreading some of them.

Here we see that models, despite their reputation for impartiality, reflect goals and ideology. When I removed the possibility of eating Pop-Tarts at every meal, I was imposing my ideology on the meals model. It's something we do without a second thought. Our own values and desires influence our choices, from the data we choose to collect to the questions we ask. Models are opinions embedded in mathematics.

Whether or not a model works is also a matter of opinion. After all, a key component of every model, whether formal or informal, is its definition of success. This is an important point that we'll return to as we explore the dark world of WMDs. In each case, we must ask not only who designed the model but also what that person or company is trying to accomplish. If the North Korean government built a model for my family's meals, for example, it might be optimized to keep us above the threshold of starvation at the lowest cost, based on the food stock available. Preferences would count for little or nothing. By contrast, if my kids were creating the model, success might feature ice cream at every meal. My own model attempts to blend a bit of the North Koreans' resource management with the happiness of my kids, along with my own priorities of health, convenience, diversity of experience, and sustainability. As a result, it's much more complex. But it still reflects my own personal reality. And a model built for today will work a bit worse tomorrow. It will grow stale if it's not constantly updated. Prices change, as do people's preferences. A model built for a six-year-old won't work for a teenager.

This is true of internal models as well. You can often see troubles when grandparents visit a grandchild they haven't seen for a while. On their previous visit, they gathered data on what the child knows, what makes her laugh, and what TV show she likes and (unconsciously) created a model for relating to this particular four-year-old. Upon meeting her a year later, they can suffer a few awkward hours because their models are out of date. Thomas the Tank Engine, it turns out, is no longer cool. It takes some time to gather new data about the child and adjust their models.

This is not to say that good models cannot be primitive. Some very effective ones hinge on a single variable. The most common model for detecting fires in a home or office weighs only one strongly correlated variable, the presence of smoke. That's usually enough. But modelers run into problems—or subject *us* to problems—when they focus models as simple as a smoke alarm on their fellow humans.

Racism, at the individual level, can be seen as a predictive model whirring away in billions of human minds around the world. It is built from faulty, incomplete, or generalized data. Whether it comes from experience or hearsay, the data indicates that certain types of people have behaved badly. That generates a binary prediction that all people of that race will behave that same way.

Needless to say, racists don't spend a lot of time hunting down reliable data to train their twisted models. And once their model morphs into a belief, it becomes hardwired. It generates poisonous assumptions, yet rarely tests them, settling instead for data that seems to confirm and fortify them. Consequently, racism is the most slovenly of predictive models. It is powered by haphazard data gathering and spurious correlations, reinforced by institutional inequities, and polluted by confirmation bias. In this way, oddly enough, racism operates like many of the WMDs I'll be describing in this book.

■ ■ ■

In 1997, a convicted murderer, an African American man named Duane Buck, stood before a jury in Harris County, Texas. Buck had killed two people, and the jury had to decide whether he would be sentenced to death or to life in prison with the chance of parole. The prosecutor pushed for the death penalty, arguing that if Buck were let free he might kill again.

Buck's defense attorney brought forth an expert witness, a psychologist named Walter Quijano, who didn't help his client's case one bit. Quijano, who had studied recidivism rates in the Texas prison system, made a reference to Buck's race, and during cross-examination the prosecutor jumped on it.

"You have determined that the … the race factor, black, increases the future dangerousness for various complicated reasons. Is that correct?" the prosecutor asked.

"Yes," Quijano answered. The prosecutor stressed that testimony in her summation, and the jury sentenced Buck to death.

Three years later, Texas attorney general John Cornyn found that the psychologist had given similar race-based testimony in six other capital cases, most of them while he worked for the prosecution. Cornyn, who would be elected in 2002 to the US Senate, ordered new race-blind hearings for the seven inmates. In a press release, he declared: "It is inappropriate to allow race to be considered as a factor in our criminal justice system. … The people of Texas want and deserve a system that affords the same fairness to everyone."

Six of the prisoners got new hearings but were again sentenced to death. Quijano's prejudicial testimony, the court ruled, had not been decisive. Buck never got a new hearing, perhaps because it was his own witness who had brought up race. He is still on death row.

Regardless of whether the issue of race comes up explicitly at trial, it has long been a major factor in sentencing. A University of Maryland study showed that in Harris County, which includes Houston, prosecutors were three times more likely to seek the death penalty for African Americans, and four times more likely for Hispanics, than for whites convicted of the same charges. That pattern isn't unique to Texas. According to the American Civil Liberties Union, sentences imposed on black men in the federal system are nearly 20 percent longer than those for whites convicted of similar crimes. And though they make up only 13 percent of the population, blacks fill up 40 percent of America's prison cells.

So you might think that computerized risk models fed by data would reduce the role of prejudice in sentencing and contribute to more even-handed treatment. With that hope, courts in twenty-four states have turned to so-called recidivism models. These help judges assess the danger posed by each convict. And by many measures they're an improvement. They keep sentences more consistent and less likely to be swayed by the moods and biases of judges. They also save money by nudging down the length of the average sentence. (It costs an average of $31,000 a year to house an inmate, and double that in expensive states like Connecticut and New York.)

The question, however, is whether we've eliminated human bias or simply camouflaged it with technology. The new recidivism models are complicated and mathematical. But embedded within these models are a host of assumptions, some of them prejudicial. And while Walter Quijano's words were transcribed for the record, which could later be read and challenged in court, the workings of a recidivism model are tucked away in algorithms, intelligible only to a tiny elite.

One of the more popular models, known as LSI–R, or Level of Service Inventory–Revised, includes a lengthy questionnaire for the prisoner to fill out. One of the questions—"How many prior convictions have you had?"—is highly relevant to the risk of recidivism. Others are also clearly related: "What part did others play in the offense? What part did drugs and alcohol play?"

But as the questions continue, delving deeper into the person's life, it's easy to imagine how inmates from a privileged background would answer one way and those from tough inner-city streets another. Ask a criminal who grew up in comfortable suburbs about "the first time you were ever involved with the police," and he might not have a single incident to report other than the one that brought him to prison. Young black males, by contrast, are likely to have been stopped by police dozens of times, even when they've done nothing wrong. A 2013 study by the New York Civil Liberties Union found that while black and Latino males between the ages of fourteen and twenty-four made up only 4.7 percent of the city's population, they accounted for 40.6 percent of the stop-and-frisk checks by police. More than 90 percent of those stopped were innocent. Some of the others might have been drinking underage or carrying a joint. And unlike most rich kids, they got in trouble for it. So if early "involvement" with the police signals recidivism, poor people and racial minorities look far riskier.

The questions hardly stop there. Prisoners are also asked about whether their friends and relatives have criminal records. Again, ask that question to a convicted criminal raised in a middle-class neighborhood, and the chances are much greater that the answer will be no. The questionnaire does avoid asking about race, which is illegal. But with the wealth of detail each prisoner provides, that single illegal question is almost superfluous.

The LSI–R questionnaire has been given to thousands of inmates since its invention in 1995. Statisticians have used those results to devise a system in which answers highly correlated to recidivism weigh more heavily and count for more points. After answering the questionnaire, convicts are categorized as high, medium, and low risk on the basis of the number of points they accumulate. In some states, such as Rhode Island, these tests are used only to target those with high-risk scores for antirecidivism programs while incarcerated. But in others, including Idaho and Colorado, judges use the scores to guide their sentencing.

This is unjust. The questionnaire includes circumstances of a criminal's birth and upbringing, including his or her family, neighborhood, and friends. These details should not be relevant to a criminal case or to the sentencing. Indeed, if a prosecutor attempted to tar a defendant by mentioning his brother's criminal record or the high crime rate in his neighborhood, a decent defense attorney would roar, "Objection, Your Honor!" And a serious judge would sustain it. This is the basis of our legal system. We are judged by what we do, not by who we are. And although we don't know the exact weights that are attached to these parts of the test, any weight above zero is unreasonable.

Many would point out that statistical systems like the LSI–R are effective in gauging recidivism risk—or at least more accurate than a judge's random guess. But even if we put aside, ever so briefly, the crucial issue of fairness, we find ourselves descending into a pernicious WMD feedback loop. A person who scores as "high risk" is likely to be unemployed and to come from a neighborhood where many of his friends and family have had run-ins with the law. Thanks in part to the resulting high score on the evaluation, he gets a longer sentence, locking him away for more years in a prison where he's surrounded by fellow criminals—which raises the likelihood that he'll return to prison. He is finally released into the same poor neighborhood, this time with a criminal record, which makes it that much harder to find a job. If he commits another crime, the recidivism model can claim another success. But in fact the model itself contributes to a toxic cycle and helps to sustain it. That's a signature quality of a WMD.

■ ■ ■

In this chapter, we've looked at three kinds of models. The baseball models, for the most part, are healthy. They are transparent and continuously updated, with both the assumptions and the conclusions clear for all to see. The models feed on statistics from the game in question, not from proxies. And the people being modeled understand the process and share the model's objective: winning the World Series. (Which isn't to say that many players, come contract time, won't quibble with a model's valuations: "Sure I struck out two hundred times, but look at my *home runs* …")

From my vantage point, there's certainly nothing wrong with the second model we discussed, the hypothetical family meal model. If my kids were to question the assumptions that underlie it, whether economic or dietary, I'd be all too happy to provide them. And even though they sometimes grouse when facing something green, they'd likely admit, if pressed, that they share the goals of convenience, economy, health, and good taste—though they might give them different weights in their own models. (And they'll be free to create them when they start buying their own food.)

I should add that my model is highly unlikely to scale. I don't see Walmart or the US Agriculture Department or any other titan embracing my app and imposing it on hundreds of millions of people, like some of the WMDs we'll be discussing. No, my model is benign, especially since it's unlikely ever to leave my head and be formalized into code.

The recidivism example at the end of the chapter, however, is a different story entirely. It gives off a familiar and noxious odor. So let's do a quick exercise in WMD taxonomy and see where it fits.

The first question: Even if the participant is aware of being modeled, or what the model is used for, is the model opaque, or even invisible? Well, most of the prisoners filling out mandatory questionnaires aren't stupid. They at least have reason to suspect that information they provide will be used against them to control them while in prison and perhaps lock them up for longer. They know the game. But prison officials know it, too. And they keep quiet about the purpose of the LSI–R questionnaire. Otherwise, they know, many prisoners will attempt to game it, providing answers to make them look like model citizens the day they leave the joint. So the prisoners are kept in the dark as much as possible and do not learn their risk scores.

In this, they're hardly alone. Opaque and invisible models are the rule, and clear ones very much the exception. We're modeled as shoppers and couch potatoes, as patients and loan applicants, and very little of this do we see—even in applications we happily sign up for. Even when such models behave themselves, opacity can lead to a feeling of unfairness. If you were told by an usher, upon entering an open-air concert, that you couldn't sit in the first ten rows of seats, you might find it unreasonable. But if it were explained to you that the first ten rows were being reserved for people in wheelchairs, then it might well make a difference. Transparency matters.

And yet many companies go out of their way to hide the results of their models or even their existence. One common justification is that the algorithm constitutes a "secret sauce" crucial to their business. It's *intellectual property,* and it must be defended, if need be, with legions of lawyers and lobbyists. In the case of web giants like Google, Amazon, and Facebook, these precisely tailored algorithms alone are worth hundreds of billions of dollars. WMDs are, by design, inscrutable black boxes. That makes it extra hard to definitively answer the second question: Does the model work against the subject's interest? In short, is it unfair? Does it damage or destroy lives?

Here, the LSI–R again easily qualifies as a WMD. The people putting it together in the 1990s no doubt saw it as a tool to bring evenhandedness and efficiency to the criminal justice system. It could also help nonthreatening criminals land lighter sentences. This would translate into more years of freedom for them and enormous savings for

American taxpayers, who are footing a $70 billion annual prison bill. However, because the questionnaire judges the prisoner by details that would not be admissible in court, it is unfair. While many may benefit from it, it leads to suffering for others.

A key component of this suffering is the pernicious feedback loop. As we've seen, sentencing models that profile a person by his or her circumstances help to create the environment that justifies their assumptions. This destructive loop goes round and round, and in the process the model becomes more and more unfair.

The third question is whether a model has the capacity to grow exponentially. As a statistician would put it, can it scale? This might sound like the nerdy quibble of a mathematician. But scale is what turns WMDs from local nuisances into tsunami forces, ones that define and delimit our lives. As we'll see, the developing WMDs in human resources, health, and banking, just to name a few, are quickly establishing broad norms that exert upon us something very close to the power of law. If a bank's model of a high-risk borrower, for example, is applied to you, the world will treat you as just that, a deadbeat—even if you're horribly misunderstood. And when that model scales, as the credit model has, it affects your whole life—whether you can get an apartment or a job or a car to get from one to the other.

When it comes to scaling, the potential for recidivism modeling continues to grow. It's already used in the majority of states, and the LSI–R is the most common tool, used in at least twenty-four of them. Beyond LSI–R, prisons host a lively and crowded market for data scientists. The penal system is teeming with data, especially since convicts enjoy even fewer privacy rights than the rest of us. What's more, the system is so miserable, overcrowded, inefficient, expensive, and inhumane that it's crying out for improvements. Who wouldn't want a cheap solution like this?

Penal reform is a rarity in today's polarized political world, an issue on which liberals and conservatives are finding common ground. In early 2015, the conservative Koch brothers, Charles and David, teamed up with a liberal think tank, the Center for American Progress, to push for prison reform and drive down the incarcerated population. But my suspicion is this: their bipartisan effort to reform prisons, along with legions of others, is almost certain to lead to the efficiency and perceived fairness of a data-fed solution. That's the age we live in.

Even if other tools supplant LSI–R as its leading WMD, the prison system is likely to be a powerful incubator for WMDs on a grand scale.

So to sum up, these are the three elements of a WMD: Opacity, Scale, and Damage. All of them will be present, to one degree or another, in the examples we'll be covering. Yes, there will be room for quibbles. You could argue, for example, that the recidivism scores are not totally opaque, since they spit out scores that prisoners, in some cases, can see. Yet they're brimming with mystery, since the prisoners cannot see how their answers produce their score. The scoring algorithm is hidden. A couple of the other WMDs might not seem to satisfy the prerequisite for scale. They're not huge, at least not yet. But they represent dangerous species that are primed to grow, perhaps exponentially. So I count them. And finally, you might note that not all of these WMDs are universally damaging. After all, they send some people to Harvard, line others up for cheap loans or good jobs, and reduce jail sentences for certain lucky felons. But the point is not whether some people benefit. It's that so many suffer. These models, powered by algorithms, slam doors in the face of millions of people, often for the flimsiest of reasons, and offer no appeal. They're unfair.

And here's one more thing about algorithms: they can leap from one field to the next, and they often do. Research in epidemiology can hold insights for box office predictions; spam filters are being retooled to identify the AIDS virus. This is true of WMDs as well. So if mathematical models in prisons appear to succeed at their job—which really boils down to efficient management of people—they could spread into the rest of the economy along with the other WMDs, leaving us as collateral damage.

That's my point. This menace is rising. And the world of finance provides a cautionary tale.

2

Shell Shocked

My Journey of Disillusionment

Imagine you have a routine. Every morning before catching the train from Joliet to Chicago's LaSalle Street station, you feed $2 into the coffee machine. It returns two quarters and a cup of coffee. But one day it returns four quarters. Three times in the next month the same machine delivers the same result. A pattern is developing.

Now, if this were a tiny anomaly in financial markets, and not a commuter train, a quant at a hedge fund—someone like me—could zero in on it. It would involve going through years of data, even decades, and then training an algorithm to predict this one recurring error—a fifty-cent swing in price—and to place bets on it. Even the smallest patterns can bring in millions to the first investor who unearths them. And they'll keep churning out profits until one of two things happens: either the phenomenon comes to an end or the rest of the market catches on to it, and the opportunity vanishes. By that point, a good quant will be hot on the trail of dozens of other tiny wrinkles.

The quest for what quants call market inefficiencies is like a treasure hunt. It can be fun. And as I got used to my new job at D. E. Shaw, I found it a welcome change from academia. While I had loved teaching at Barnard, and had loved my research on algebraic number theory, I found progress agonizingly slow. I wanted to be part of the fast-paced real world.

At that point, I considered hedge funds morally neutral—scavengers in the financial system, at worst. I was proud to go to Shaw, known as the Harvard of the hedge funds, and show the people there that my smarts could translate into money. Plus, I would be earning three times what I had earned as a professor. I could hardly suspect, as I began my new job, that it would give me a front-row seat during the financial crisis and a terrifying tutorial on how insidious and destructive math could be. At the hedge fund, I got my first up-close look at a WMD.

In the beginning, there was plenty to like. Everything at Shaw was powered by math. At a lot of firms, the traders run the show, making big deals, barking out orders, and landing multimillion-dollar bonuses. Quants are their underlings. But at Shaw the traders are little more than functionaries. They're called executioners. And the mathematicians reign supreme. My ten-person team was the "futures group." In a business in which everything hinges on what will happen tomorrow, what could be bigger than that?

We had about fifty quants in total. In the early days, it was entirely men, except for me. Most of them were foreign born. Many of them had come from abstract math or physics; a few, like me, had come from number theory. I didn't get much of a chance to talk shop with them, though. Since our ideas and algorithms were the foundation of the hedge fund's business, it was clear that we quants also represented a risk: if we walked away, we could quickly use our knowledge to fuel a fierce competitor.

To keep this from happening on a large, firm-threatening scale, Shaw mostly prohibited us from talking to colleagues in other groups—or sometimes even our own office mates—about what we were doing. In a sense, information was cloistered in a networked cell structure, not unlike that of Al Qaeda. That way, if one cell collapsed—if one of us hightailed it to Bridgewater or J.P. Morgan, or set off on our own—we'd take with us only our own knowledge. The rest of Shaw's business would carry on unaffected. As you can imagine, this wasn't terrific for camaraderie.

Newcomers were required to be on call every thirteen weeks in the futures group. This meant being ready to respond to computer problems whenever any of the world's markets were open, from Sunday evening our time, when the Asian markets came to life, to New York's closing bell at 4 p.m. on Friday. Sleep deprivation was an issue. But worse was the powerlessness to respond to issues in a shop that didn't share information. Say an algorithm appeared to be misbehaving. I'd have to locate it and then find the person responsible for it, at any time of the day or night, and tell him (and it was always a him) to fix it. It wasn't always a friendly encounter.

Then there were panics. Over holidays, when few people were working, weird things tended to happen. We had all sorts of things in our huge portfolio, including currency forwards, which were promises to buy large amounts of a foreign currency in a couple of days. Instead of actually buying the foreign currency, though, a trader would "roll over" the position each day so the promise would be put off for one more day. This way, our bet on the direction of the market would be sustained but we'd never have to come up with loads of cash. One time over Christmas I noticed a large position in Japanese yen that was coming due. Someone had to roll that contract over. This was a job typically handled by a colleague in Europe, who presumably was home with his family. I saw that if it didn't happen soon someone theoretically would have to show up in Tokyo with $50 million in yen. Ironing out that problem added a few frantic hours to the holiday.

All of those issues might fit into the category of occupational hazard. But the real problem came from a nasty feeling I started to have in my stomach. I had grown accustomed to playing in these oceans of currency, bonds, and equities, the trillions of dollars flowing through international markets. But unlike the numbers in my academic models, the figures in my models at the hedge fund stood for something. They were people's retirement funds and mortgages. In retrospect, this seems blindingly obvious. And of course, I knew it all along, but I hadn't truly appreciated the nature of the nickels, dimes, and quarters that we pried loose with our mathematical tools. It wasn't found money, like nuggets from a mine or coins from a sunken Spanish galleon. This wealth was coming out of people's pockets. For hedge funds, the smuggest of the players on Wall Street, this was "dumb money."

It was when the markets collapsed in 2008 that the ugly truth struck home in a big way. Even worse than filching dumb money from people's accounts, the finance industry was in the business of creating WMDs, and I was playing a small part.

The troubles had actually started a year earlier. In July of 2007, "interbank" interest rates spiked. After the recession that followed the terrorist attacks in 2001, low interest rates had fueled a housing boom. Anyone, it seemed, could get a mortgage, builders were turning exurbs, desert, and prairie into vast new housing developments, and banks gambled billions on all kinds of financial instruments tied to the building bonanza.

But these rising interest rates signaled trouble. Banks were losing trust in each other to pay back overnight loans. They were slowly coming to grips with the dangerous junk they held in their own portfolios and judged, wisely, that others were sitting on just as much risk, if not more. Looking back, you could say the interest rate spikes were actually a sign of sanity, although they obviously came too late.

At Shaw, these jitters dampened the mood a bit. Lots of companies were going to struggle, it was clear. The industry was going to take a hit, perhaps a very big one. But still, it might not be our problem. We didn't plunge headlong into risky markets. Hedge funds, after all, hedged. That was our nature. Early on, we called the market turbulence "the kerfuffle." For Shaw, it might cause some discomfort, maybe even an embarrassing episode or two, like when a rich man's credit card is denied at a fancy restaurant. But there was a good chance we'd be okay.

Hedge funds, after all, didn't make these markets. They just played in them. That meant that when the market crashed, as it would, rich opportunities would emerge from the wreckage. The game for hedge funds was not so much to ride markets up as to predict the movements within them. Down could be every bit as lucrative.

To understand how hedge funds operate at the margins, picture a World Series game at Chicago's Wrigley Field. With a dramatic home run in the bottom of the ninth inning, the Cubs win their first championship since 1908, back when Teddy Roosevelt was president. The stadium explodes in celebration. But a single row of fans stays seated, quietly analyzing a slew of results. These gamblers don't hold the traditional win-or-lose bets. Instead they may have bet that Yankees relievers would give up more walks than strikeouts, that the game would feature at least one bunt but no more than two, or that the Cubs' starter would last at least six innings. They even hold bets that other gamblers will win or lose their own bets. These people wager on many movements associated with the game, but not as much on the game itself. In this, they behave like hedge funds.

That made us feel safe, or at least safer. I remember a gala event to celebrate the architects of the system that would soon crash. The firm welcomed Alan Greenspan, the former Fed chairman, and Robert Rubin, the former Treasury secretary and Goldman Sachs executive. Rubin had pushed for a 1999 revision of the Depression-era Glass-Steagall Act. This removed the glass wall between banking and investment operations, which facilitated the orgy of speculation over the following decade. Banks were free to originate loans (many of them fraudulent) and sell them to their customers in the form of securities.

That wasn't so unusual and could be considered a service they did for their customers. However, now that Glass-Steagall was gone, the banks could, and sometimes did, bet against the very same securities that they'd sold to customers. This created mountains of risk—and endless investment potential for hedge funds. We placed our bets, after all, on market movements, up or down, and those markets were frenetic.

At the D. E. Shaw event, Greenspan warned us about problems in mortgage-backed securities. That memory nagged me when I realized a couple of years later that Rubin, who at the time worked at Citigroup, had been instrumental in collecting a massive portfolio of these exact toxic contracts—a major reason Citigroup later had to be bailed out at taxpayer expense.

Sitting with these two was Rubin's protégé and our part-time partner, Larry Summers. He had followed Rubin in Treasury and had gone on to serve as president of Harvard University. Summers had troubles with faculty, though. And professors had risen up against him in part because he suggested that the low numbers of women in math and the hard sciences might be due to genetic inferiority—what he called the unequal distribution of "intrinsic aptitude."

After Summers left the Harvard presidency, he landed at Shaw. And I remember that when it came time for our founder, David Shaw, to address the prestigious trio, he joked that Summers's move from Harvard to Shaw had been a "promotion." The markets might be rumbling, but Shaw was still on top of the world.

Yet as the crisis deepened the partners at Shaw lost a bit of their swagger. Troubled markets, after all, were entwined. For example, rumors were already circulating about the vulnerability of Lehman Brothers, which owned 20 percent of D. E. Shaw and handled many of our transactions. As the markets continued to rattle and shake, the internal mood turned fretful. We could crunch numbers with the best of the best. But what if the frightening tomorrow on the horizon didn't resemble any of the yesterdays? What if it was something entirely new and different?

That was a concern, because mathematical models, by their nature, are based on the past, and on the assumption that patterns will repeat. Before long, the equities group liquidated its holdings, at substantial cost. And the hiring spree for new quants, which had brought me to the firm, ended. Although people tried to laugh off this new climate, there was a growing fear. All eyes were on securitized products, especially the mortgage-backed securities Greenspan had warned us about.

For decades, mortgage securities had been the opposite of scary. They were boring financial instruments that individuals and investment funds alike used to diversify their portfolios. The idea behind them was that quantity could offset risk. Each single mortgage held potential for default: the home owner could declare bankruptcy, meaning the bank would never be able to recover all of the money it had loaned. At the other extreme, the borrower could pay back the mortgage ahead of schedule, bringing the flow of interest payments to a halt.

And so in the 1980s, investment bankers started to buy thousands of mortgages and package them into securities—a kind of bond, which is to say an instrument that pays regular dividends, often at quarterly intervals. A few of the home owners would default, of course. But most people would stay afloat and keep paying their mortgages, generating a smooth and predictable flow of revenue. In time, these bonds grew into an entire industry, a pillar of the capital markets. Experts grouped the mortgages into different classes, or tranches. Some were considered rock solid. Others carried more risk—and higher interest rates. Investors had reason to feel confident because the credit-rating agencies, Standard & Poor's, Moody's, and Fitch, had studied the securities and scored them for risk. They considered them sensible investments. But consider the opacity. Investors remained blind to the quality of the mortgages in the securities. Their only glimpse of what lurked inside came from analyst ratings. And these analysts collected fees from the very companies whose products they were rating. Mortgage-backed securities, needless to say, were an ideal platform for fraud.

If you want a metaphor, one commonly used in this field comes from sausages. Think of the mortgages as little pieces of meat of varying quality, and think of the mortgage-backed securities as bundles of the sausage that result from throwing everything together and adding a bunch of strong spices. Of course, sausages can vary in quality, and it's hard to tell from the outside what went into them, but since they have a stamp from the USDA saying they're safe to eat, our worries are put aside.

As the world later learned, mortgage companies were making rich profits during the boom by loaning money to people for homes they couldn't afford. The strategy was simply to write unsustainable mortgages, snarf up the fees, and then unload the resulting securities—the sausages—into the booming mortgage security market. In one notorious case, a strawberry picker named Alberto Ramirez, who made $14,000 a year, managed to finance a $720,000 house in Rancho Grande, California. His broker apparently told him that he could refinance in a few months and later flip the house and make a tidy profit. Months later, he defaulted on the loan.

In the run-up to the housing collapse, mortgage banks were not only offering unsustainable deals but actively prospecting for victims in poor and minority neighborhoods. In a federal lawsuit, Baltimore officials charged Wells Fargo with targeting black neighborhoods for so-called ghetto loans. The bank's "emerging markets" unit, according to a former bank loan officer, Beth Jacobson, focused on black churches. The idea was that trusted pastors would steer their congregants toward loans. These turned out to be subprime loans carrying the highest interest rates. The bank sold these even to borrowers with rock-solid credit, who should have qualified for loans with far better terms. By the time Baltimore filed the suit, in 2009, more than half of the properties subject to foreclosure on Well Fargo loans were empty, and 71 percent of them were in largely African American neighborhoods. (In 2012, Wells Fargo settled the suit, agreeing to pay $175 million to thirty thousand victims around the country.)

To be clear, the subprime mortgages that piled up during the housing boom, whether held by strawberry pickers in California or struggling black congregants in Baltimore, were not WMDs. They were financial instruments, not models, and they had little to do with math. (In fact, the brokers went to great lengths to ignore inconvenient numbers.)

But when banks started loading mortgages like Alberto Ramirez's into classes of securities and selling them, they were relying on flawed mathematical models to do it. The risk model attached to mortgage-backed securities was a WMD. The banks were aware that some of the mortgages were sure to default. But banks held on to two false assumptions, which sustained their confidence in the system.

The first false assumption was that crack mathematicians in all of these companies were crunching the numbers and ever so carefully balancing the risk. The bonds were marketed as products whose risk was assessed by specialists using cutting-edge algorithms. Unfortunately, this just wasn't the case. As with so many WMDs, the math was directed against the consumer as a smoke screen. Its purpose was only to optimize short-term profits for the sellers. And those sellers trusted that they'd manage to unload the securities before they exploded. Smart people would win. And dumber people, the providers of dumb money, would wind up holding billions (or trillions) of unpayable IOUs. Even rigorous mathematicians—and there were a few—were working with numbers provided by people carrying out wide-scale fraud. Very few people had the expertise and the information required to know what was actually going on statistically, and most of the people who did lacked the integrity to speak up. The risk ratings on the securities were designed to be opaque and mathematically intimidating, in part so that buyers wouldn't perceive the true level of risk associated with the contracts they owned.

The second false assumption was that not many people would default at the same time. This was based on the theory, soon to be disproven, that defaults were largely random and unrelated events. This led to a belief that solid mortgages would offset the losers in each tranche. The risk models were assuming that the future would be no different from the past.

In order to sell these mortgage-backed bonds, the banks needed AAA ratings. For this, they looked to the three credit-rating agencies. As the market expanded, rating the growing billion-dollar market in mortgage bonds turned into a big business for the agencies, bringing in lucrative fees. They grew addicted to those fees. And they understood all too clearly that if they provided anything less than AAA ratings, the banks would take the work to their competitors. So the agencies played ball. They paid more attention to customer satisfaction than to the accuracy of their models. These risk models also created their own pernicious feedback loop. The AAA ratings on defective products turned into dollars. The dollars in turn created confidence in the products and in the cheating-and-lying process that manufactured them. The resulting cycle of mutual back-scratching and pocket-filling was how the whole sordid business operated until it blew up.

Of all the WMD qualities, the one that turned these risk models into a monstrous force of global dimension was scale. Snake oil vendors, of course, are as old as history, and in previous real estate bubbles unwitting buyers ended up with swampland and stacks of false deeds. But this time the power of modern computing fueled fraud at a scale unequaled in history. The damage was compounded by other vast markets that had grown up around the mortgage-backed securities:

credit default swaps and synthetic collateralized debt obligations, or CDOs. Credit default swaps were small insurance policies that transferred the risk on a bond. The swaps gave banks and hedge funds alike a sense of security, since they could supposedly use them to balance risk. But if the entities holding these insurance policies go belly up, as many did, the chain reaction blows holes through the global economy. Synthetic CDOs went one step further: they were contracts whose value depended on the performance of credit default swaps and mortgage-backed securities. They allowed financial engineers to leverage up their bets even more.

The overheated (and then collapsing) market featured $3 trillion of subprime mortgages by 2007, and the market around it—including the credit default swaps and synthetic CDOs, which magnified the risks—was twenty times as big. No national economy could compare.

Paradoxically, the supposedly powerful algorithms that created the market, the ones that analyzed the risk in tranches of debt and sorted them into securities, turned out to be useless when it came time to clean up the mess and calculate what all the paper was actually worth. The math could multiply the horseshit, but it could not decipher it. This was a job for human beings. Only people could sift through the mortgages, picking out the false promises and wishful thinking and putting real dollar values on the loans. It was a painstaking process, because people—unlike WMDs—cannot scale their work exponentially, and for much of the industry it was a low priority. During this lengthy detox, of course, the value of the debt—and the homes that the debt relied on—kept falling. And as the economy took a nosedive, even home owners who could afford their mortgages when the crisis began were suddenly at risk of defaulting, too.

As I've mentioned, Shaw was a step or two removed from the epicenter of the market collapse. But as other players started to go under, they were frantically undoing trades that affected the ones we had on our books. It had a cascading effect, and as we entered the second half of 2008 we were losing money left and right.

Over the following months, disaster finally hit the mainstream. That's when everyone finally saw the people on the other side of the algorithms. They were desperate home owners losing their homes and millions of Americans losing their jobs. Credit card defaults leapt to record highs. The human suffering, which had been hidden from view behind numbers, spreadsheets, and risk scores, became palpable.

The chatter at Shaw was nervous. After the fall of Lehman Brothers in September of 2008, people discussed the political fallout. Barack Obama looked likely to win the election in November. Would he hammer the industry with new regulations? Raise taxes on carried interest? These people weren't losing their houses or maxing out their credit cards just to stay afloat. But they found plenty to worry about, just the same. The only choice was to wait it out, let the lobbyists do their work, and see if we'd be allowed to continue as usual.

By 2009, it was clear that the lessons of the market collapse had brought no new direction to the world of finance and had instilled no new values. The lobbyists succeeded, for the most part, and the game remained the same: to rope in dumb money. Except for a few regulations that added a few hoops to jump through, life went on.

This drama pushed me quickly along in my journey of disillusionment. I was especially disappointed in the part that mathematics had played. I was forced to confront the ugly truth: people had deliberately wielded formulas to impress rather than clarify. It was the first time I had been directly confronted with this toxic concept, and it made me want to escape, to go back in time to the world of proofs and Rubik's Cubes.

And so I left the hedge fund in 2009 with the conviction that I would work to fix the financial WMDs. New regulations were forcing banks to hire independent experts to analyze their risk. I went to work for one of the companies providing that analysis, RiskMetrics Group, one block north of Wall Street. Our product was a blizzard of numbers, each of them predicting the likelihood that a certain tranche of securities or commodities would go poof within the next week, the next year, or the next five years. When everyone is betting on everything that moves in the market, a smart read on risk is worth gold.

To calculate risk, our team employed the Monte Carlo method. To picture it, just imagine spinning the roulette wheel at a casino ten thousand times, taking careful notes all the while. Using Monte Carlo, you'd typically start with historical market data and run through thousands of test scenarios. How would the portfolio we're studying fare on each trading day since 2010, or 2005? Would it survive the very darkest days of the crash? How likely is it that a mortal threat will arise in the next year or two? To come up with these odds, scientists run thousands upon thousands of simulations. There was plenty to complain about with this method, but it was a simple way to get some handle on your risk.

My job was to act as a liaison between our risk management business and the largest and most discerning connoisseurs of risk, the quantitative hedge funds. I'd call the hedge funds, or they'd call me, and we'd discuss any questions they had about our numbers. As often as not, though, they'd notify me only when we'd made a mistake. The fact was, the hedge funds always considered themselves the smartest of the smart, and since understanding risk was fundamental to their existence, they would never rely entirely on outsiders like us. They had their own risk teams, and they bought our product mostly to look good for investors.

I also answered the hotline and would sometimes find myself answering questions from clients at big banks. Eager to repair their tattered image, they wanted to be viewed as responsible, which is why they were calling in the first place. But, unlike the hedge funds, they showed little interest in our analysis. The risk in their portfolios was something they almost seemed to ignore. Throughout my time at the hotline, I got the sense that the people warning about risk were viewed as party poopers or, worse, a threat to the bank's bottom line. This was true even after the cataclysmic crash of 2008, and it's not hard to understand why. If they survived that one—because they were too big to fail—why were they going to fret over risk in their portfolio now?

The refusal to acknowledge risk runs deep in finance. The culture of Wall Street is defined by its traders, and risk is something they actively seek to underestimate. This is a result of the way we define a trader's prowess, namely by his "Sharpe ratio," which is calculated as the profits he generates divided by the risks in his portfolio. This ratio is crucial to a trader's career, his annual bonus, his very sense of being. If you disembody those traders and consider them as a set of algorithms, those algorithms are relentlessly focused on optimizing the Sharpe ratio. Ideally, it will climb, or at least never fall too low. So if one of the risk reports on credit default swaps bumped up the risk calculation on one of a trader's key holdings, his Sharpe ratio would tumble. This could cost him hundreds of thousands of dollars when it came time to calculate his year-end bonus.

I soon realized that I was in the rubber-stamp business. In 2011 it was time to move again, and I saw a huge growth market for mathematicians like me. In the time it took me to type two words into my résumé, I was a newly proclaimed Data Scientist, and ready to plunge into the Internet economy. I landed a job at a New York start-up called Intent Media.

I started out building models to anticipate the behavior of visitors to various travel websites. The key question was whether someone showing up at the Expedia site was just browsing or looking to spend money. Those who weren't planning to buy were worth very little in potential revenue. So we would show them comparison ads for competing services such as Travelocity or Orbitz. If they clicked on the ad, it brought in a few pennies, which was better than nothing. However, we didn't want to feed these ads to serious shoppers. In the worst case, we'd gain a dime of ad revenue while sending potential customers to rivals, where perhaps they'd spend thousands of dollars on hotel rooms in London or Tokyo. It would take thousands of ad views to make up for even a few hundred dollars in lost fees. So it was crucial to keep those people in house.

My challenge was to design an algorithm that would distinguish window shoppers from buyers. There were a few obvious signals. Were they logged into the service? Had they bought there before? But I also scoured for other hints. What time of day was it, and what day of the year? Certain weeks are hot for buyers. The Memorial Day "bump," for example, occurs in mid-spring, when large numbers of people make summer plans almost in unison. My algorithm would place a higher value on shoppers during these periods, since they were more likely to buy.

The statistical work, as it turned out, was highly transferable from the hedge fund to e-commerce—the biggest difference was that, rather than the movement of markets, I was now predicting people's clicks.

In fact, I saw all kinds of parallels between finance and Big Data. Both industries gobble up the same pool of talent, much of it from elite universities like MIT, Princeton, or Stanford. These new hires are ravenous for success and have been focused on external metrics—like SAT scores and college admissions—their entire lives. Whether in finance or tech, the message they've received is that they will be rich, that they will run the world. Their productivity indicates that they're on the right track, and it translates into dollars. This leads to the fallacious conclusion that whatever they're doing to bring in more money is good. It "adds value." Otherwise, why would the market reward it?

In both cultures, wealth is no longer a means to get by. It becomes directly tied to personal worth. A young suburbanite with every advantage—the prep school education, the exhaustive coaching for college admissions tests, the overseas semester in Paris or Shanghai—still flatters himself that it is his skill, hard work, and prodigious problem-solving abilities that have lifted him into a world of privilege. Money vindicates all doubts. And the rest of his circle plays along, forming a mutual admiration society. They're eager to convince us all that Darwinism is at work, when it looks very much to the outside like a combination of gaming a system and dumb luck.

In both of these industries, the real world, with all of its messiness, sits apart. The inclination is to replace people with data trails, turning them into more effective shoppers, voters, or workers to optimize some objective. This is easy to do, and to justify, when success comes back as an anonymous score and when the people affected remain every bit as abstract as the numbers dancing across the screen.

I was already blogging as I worked in data science, and I was also getting more involved with the Occupy movement. More and more, I worried about the separation between technical models and real people, and about the moral repercussions of that separation. In fact, I saw the same pattern emerging that I'd witnessed in finance: a false sense of security was leading to widespread use of imperfect models, self-serving definitions of success, and growing feedback loops. Those who objected were regarded as nostalgic Luddites.

I wondered what the analogue to the credit crisis might be in Big Data. Instead of a bust, I saw a growing dystopia, with inequality rising. The algorithms would make sure that those deemed losers would remain that way. A lucky minority would gain ever more control over the data economy, raking in outrageous fortunes and convincing themselves all the while that they deserved it.

After a couple of years working and learning in the Big Data space, my journey to disillusionment was more or less complete, and the misuse of mathematics was accelerating. In spite of blogging almost daily, I could barely keep up with all the ways I was hearing of people being manipulated, controlled, and intimidated by algorithms. It started with teachers I knew struggling under the yoke of the value-added model, but it didn't end there. Truly alarmed, I quit my job to investigate the issue in earnest.

10

The Targeted Citizen

Civic Life

As you know by now, I am outraged by all sorts of WMDs. So let's imagine that I decide to launch a campaign for tougher regulations on them, and I post a petition on my Facebook page. Which of my friends will see it on their news feed?

I have no idea. As soon as I hit send, that petition belongs to Facebook, and the social network's algorithm makes a judgment about how to best use it. It calculates the odds that it will appeal to each of my friends. Some of them, it knows, often sign petitions, and perhaps share them with their own networks. Others tend to scroll right past. At the same time, a number of my friends pay more attention to me and tend to click the articles I post. The Facebook algorithm takes all of this into account as it decides who will see my petition. For many of my friends, it will be buried so low on their news feed that they'll never see it.

This is what happens when the immensely powerful network we share with 1.5 billion users is also a publicly traded corporation. While Facebook may feel like a modern town square, the company determines, according to its own interests, what we see and learn on its social network. As I write this, about two-thirds of American adults have a profile on Facebook. They spend thirty-nine minutes a day on the site, only four minutes less than they dedicate to face-to-face socializing. Nearly half of them, according to a Pew Research Center report, count on Facebook to deliver at least some of their news, which leads to the question: By tweaking its algorithm and molding the news we see, can Facebook game the political system?

The company's own researchers have been looking into this. During the 2010 and 2012 elections, Facebook conducted experiments to hone a tool they called the "voter megaphone." The idea was to encourage people to spread word that they had voted. This seemed reasonable enough. By sprinkling people's news feeds with "I voted" updates, Facebook was encouraging Americans—more than sixty-one million of them—to carry out their civic duty and make their voices heard. What's more, by posting about people's voting behavior, the site was stoking peer pressure to vote. Studies have shown that the quiet satisfaction of carrying out a civic duty is less likely to move people than the possible judgment of friends and neighbors.

At the same time, Facebook researchers were studying how different types of updates influenced people's voting behavior. No researcher had ever worked in a human laboratory of this scale. Within hours, Facebook could harvest information from tens of millions of people, or more, measuring the impact that their words and shared links had on each other. And it could use that knowledge to influence people's actions, which in this case happened to be voting.

That's a significant amount of power. And Facebook is not the only company to wield it. Other publicly held corporations, including Google, Apple, Microsoft, Amazon, and cell phone providers like Verizon and AT&T, have vast information on much of humanity—and the means to steer us in any way they choose.

Usually, as we've seen, they're focused on making money. However, their profits are tightly linked to government policies. The government regulates them, or chooses not to, approves or blocks their mergers and acquisitions, and sets their tax policies (often turning a blind eye to the billions parked in offshore tax havens). This is why tech companies, like the rest of corporate America, inundate Washington with lobbyists and quietly pour hundreds of millions of dollars in contributions into the political system. Now they're gaining the wherewithal to fine-tune our political behavior—and with it the shape of American government—just by tweaking their algorithms.

The Facebook campaign started out with a constructive and seemingly innocent goal: to encourage people to vote. And it succeeded. After comparing voting records, researchers estimated that their campaign had increased turnout by 340,000 people. That's a big enough crowd to swing entire states, and even national elections. George W. Bush, after all, won in 2000 by a margin of 537 votes in Florida. The activity of a single Facebook algorithm on Election Day, it's clear, could not only change the balance of Congress but also decide the presidency.

Facebook's potency comes not only from its reach but also from its ability to use its own customers to influence their friends. The vast majority of the sixty-one million people in the experiment received a message on their news feed encouraging them to vote. The message included a display of photos: six of the user's Facebook friends, randomly selected, who had clicked the "I Voted" button. The researchers also studied two control groups, each numbering around six hundred thousand. One group saw the "I Voted" campaign, but without the pictures of friends. The other received nothing at all.

By sprinkling its messages through the network, Facebook was studying the impact of friends' behavior on our own. Would people encourage their friends to vote, and would this affect their behavior? According to the researchers' calculations, seeing that friends were participating made all the difference. People paid much more attention when the "I Voted" updates came from friends, and they were more likely to share those updates. About 20 percent of the people who saw that their friends had voted also clicked on the "I Voted" button. Among those who didn't get the button from friends, only 18 percent did. We can't be sure that all the people who clicked the button actually voted, or that those who didn't click it stayed home. Still, with sixty-one million potential voters on the network, a possible difference of two points can be huge.

Two years later, Facebook took a step further. For three months leading up to the election between President Obama and Mitt Romney, a researcher at the company, Solomon Messing, altered the news feed algorithm for about two million people, all of them politically engaged. These people got a higher proportion of hard news, as opposed to the usual cat videos, graduation announcements, or photos from Disney World. If their friends shared a news story, it showed up high on their feed.

Messing wanted to see if getting more news from friends changed people's political behavior. Following the election, Messing sent out surveys. The self-reported results indicated that the voter participation in this group inched up from 64 to 67 percent. "When your friends deliver the newspaper," said Lada Adamic, a computational social scientist at Facebook, "interesting things happen." Of course, it wasn't really the friends delivering the newspaper, but Facebook itself. You might argue that newspapers have exerted similar power for eons. Editors pick the front-page news and decide how to characterize it.

They choose whether to feature bombed Palestinians or mourning Israelis, a policeman rescuing a baby or battering a protester. These choices can no doubt influence both public opinion and elections. The same goes for television news. But when the *New York Times* or CNN covers a story, everyone sees it. Their editorial decision is clear, on the record. It is not opaque. And people later debate (often on Facebook) whether that decision was the right one.

Facebook is more like the Wizard of Oz: we do not see the human beings involved. When we visit the site, we scroll through updates from our friends. The machine appears to be only a neutral go-between. Many people still believe it is. In 2013, when a University of Illinois researcher named Karrie Karahalios carried out a survey on Facebook's algorithm, she found that 62 percent of the people were unaware that the company tinkered with the news feed. They believed that the system instantly shared everything they posted with all of their friends.

The potential for Facebook to hold sway over our politics extends beyond its placement of news and its Get Out the Vote campaigns. In 2012, researchers experimented on 680,000 Facebook users to see if the updates in their news feeds could affect their mood. It was already clear from laboratory experiments that moods are contagious. Being around a grump is likely to turn you into one, if only briefly. But would such contagions spread online?

Using linguistic software, Facebook sorted positive (stoked!) and negative (bummed!) updates. They then reduced the volume of downbeat postings in half of the news feeds, while reducing the cheerful quotient in the others. When they studied the users' subsequent posting behavior, they found evidence that the doctored new feeds had indeed altered their moods. Those who had seen fewer cheerful updates produced more negative posts. A similar pattern emerged on the positive side.

Their conclusion: "Emotional states can be transferred to others …, leading people to experience the same emotions without their awareness." In other words, Facebook's algorithms can affect how millions of people feel, and those people won't know that it's happening. What would occur if they played with people's emotions on Election Day?

I have no reason to believe that the social scientists at Facebook are actively gaming the political system. Most of them are serious academics carrying out research on a platform that they could only have dreamed about two decades ago. But what they have demonstrated is Facebook's enormous power to affect what we learn, how we feel, and whether we vote. Its platform is massive, powerful, and opaque. The algorithms are hidden from us, and we see only the results of the experiments researchers choose to publish.

Much the same is true of Google. Its search algorithm appears to be focused on raising revenue. But search results, if Google so chose, could have a dramatic effect on what people learn and how they vote. Two researchers, Robert Epstein and Ronald E. Robertson, recently asked undecided voters in both the United States and India to use a search engine to learn about upcoming elections. The engines they used were programmed to skew the search results, favoring one party over another. Those results, they said, shifted voting preferences by 20 percent.

This effect was powerful, in part, because people widely trust search engines. Some 73 percent of Americans, according to a Pew Research report, believe that search results are both accurate and impartial. So companies like Google would be risking their own reputation, and inviting a regulatory crackdown, if they doctored results to favor one political outcome over another.

Then again, how would anyone know? What we learn about these Internet giants comes mostly from the tiny proportion of their research that they share. Their algorithms represent vital trade secrets. They carry out their business in the dark.

I wouldn't yet call Facebook or Google's algorithms political WMDs, because I have no evidence that the companies are using their networks to cause harm. Still, the potential for abuse is vast. The drama occurs in code and behind imposing firewalls. And as we'll see, these technologies can place each of us into our own cozy political nook.

■ ■ ■

By late spring of 2012, the former governor of Massachusetts, Mitt Romney, had the Republican nomination sewn up. The next step was to build up his war chest for the general election showdown with President Obama. And so on May 17, he traveled to Boca Raton, Florida, for a fund-raiser at the palatial home of Marc Leder, a private equity investor. Leder had already poured $225,000 into the pro-

Romney Super PAC Restore Our Future and had given another $63,330 to the Romney Victory PAC. He had gathered a host of rich friends, most of them in finance and real estate, to meet the candidate. Naturally, the affair would be catered.

Romney could safely assume that he was walking into a closed setting with a group of people who thought much like Marc Leder. If this had been a televised speech, Romney would have taken great care not to ruffle potential Republican voters. Those ranged from Evangelical Christians and Wall Street financiers to Cuban Americans and suburban soccer moms. Trying to please everyone is one reason most political speeches are boring (and Romney's, even his supporters groused, were especially so). But at an intimate gathering at Marc Leder's house, a small and influential group might get closer to the real Mitt Romney and hear what the candidate really believed, unfiltered. They had already given him large donations. A frank chat was the least they could expect for their investment.

Basking in the company of people he believed to be supportive and like-minded, Romney let loose with his observation that 47 percent of the population were "takers," living off the largesse of big government. These people would never vote for him, the governor said—which made it especially important to reach out to the other 53 percent. But Romney's targeting, it turned out, was inexact. The caterers circulating among the donors, serving drinks and canapés, were outsiders. And like nearly everyone in the developed world, they carried phones equipped with video cameras. Romney's dismissive remarks, captured by a bartender, went viral. The gaffe very likely cost Romney any chance he had of winning the White House.

Success for Romney at that Boca Raton gathering required both accurate targeting and secrecy. He wanted to be the ideal candidate for Marc Leder and friends. And he trusted that Leder's house represented a safe zone in which to be that candidate. In a dream world, politicians would navigate countless such targeted safe zones so that they could tailor their pitch for every subgroup—without letting the others see it. One candidate could be many candidates, with each part of the electorate seeing only the parts they liked.

This duplicity, or "multiplicity," is nothing new in politics. Politicians have long tried to be many things to many people, whether they're eating kielbasa in Milwaukee, quoting the Torah in Brooklyn, or pledging allegiance to corn-based ethanol in Iowa. But as Romney discovered, video cameras can now bust them if they overdo their contortions.

Modern consumer marketing, however, provides politicians with new pathways to specific voters so that they can tell them what they know they want to hear. Once they do, those voters are likely to accept the information at face value because it confirms their previous beliefs, a phenomenon psychologists call confirmation bias. It is one reason that none of the invited donors at the Romney event questioned his assertion that nearly half of voters were hungry for government handouts. It only bolstered their existing beliefs.

This merging of politics and consumer marketing has been developing for the last half century, as the tribal rituals of American politics, with their ward bosses and long phone lists, have given way to marketing science. In *The Selling of the President,* which followed Richard Nixon's 1968 campaign, the journalist Joe McGinniss introduced readers to the political operatives working to market the presidential candidate like a consumer good. By using focus groups, Nixon's campaign was able to hone his pitch for different regions and demographics.

But as time went on, politicians wanted a more detailed approach, one that would ideally reach each voter with a personalized come-on. This desire gave birth to direct-mail campaigns. Borrowing tactics from the credit card industry, political operatives built up huge databases of customers—voters, in this case—and placed them into various subgroups, reflecting their values and their demographics. For the first time, it was possible for next-door neighbors to receive different letters or brochures from the same politician, one vowing to protect wilderness and the other stressing law and order.

Direct mail was microtargeting on training wheels. The convergence of Big Data and consumer marketing now provides politicians with far more powerful tools. They can target microgroups of citizens for both votes and money and appeal to each of them with a meticulously honed message, one that no one else is likely to see. It might be a banner on Facebook or a fund-raising email. But each one allows candidates to quietly sell multiple versions of themselves—and it's anyone's guess which version will show up for work after inauguration.

■ ■ ■

In July of 2011, more than a year before President Obama would run for reelection, a data scientist named Rayid Ghani posted an update on LinkedIn:

> *Hiring analytics experts who want to make a difference. The Obama re-election campaign is growing the analytics team to work on high-impact large-scale data mining problems.*
>
> *We have several positions available at all levels of experience. Looking for experts in statistics, machine learning, data mining, text analytics, and predictive analytics to work with large amounts of data and help guide election strategy.*

Ghani, a computer scientist educated at Carnegie Mellon, would be heading up the data team for Obama's campaign. In his previous position, at Accenture Labs in Chicago, Ghani had developed consumer applications for Big Data, and he trusted that he could apply his skills to politics. The goal for the Obama campaign was to create tribes of like-minded voters, people as uniform in their values and priorities as the guests at Marc Leder's reception—but without the caterers. Then they could target them with the messaging most likely to move them toward specific objectives, including voting, organizing, and fund-raising.

One of Ghani's projects at Accenture involved modeling supermarket shoppers. A major grocer had provided the Accenture team with a massive database of anonymized consumer purchases. The idea was to dig into this data to study each consumer's buying habits and then to place the shoppers into hundreds of different consumer buckets. There would be the impulse shoppers who bought candy at the checkout counter and the health nuts who were willing to pay triple for organic kale. Those were the obvious categories. But others were more surprising. Ghani and his team, for example, could spot people who stuck close to a brand and others who would switch for even a tiny discount. There were buckets for these "persuadables," too. The end goal was to come up with a different plan for each shopper and to guide them through the store, leading them to all the foods they were most likely to want and buy.

Unfortunately for Accenture's clients, this ultimate vision hinged upon the advent of computerized shopping carts, which haven't yet caught on in a big way and maybe never will. But despite the disappointment in supermarkets, Ghani's science translated perfectly into politics. Those fickle shoppers who switched brands to save a few cents, for example, behaved very much like swing voters. In the supermarket, it was possible to estimate how much it would cost to turn each shopper from one brand of ketchup or coffee to another more profitable brand. The supermarket could then pick out, say, the 15 percent most likely to switch and provide them with coupons. Smart targeting was essential. They certainly didn't want to give coupons to shoppers who were ready to pay full price. That was like burning money.*

Would similar calculations work for swing voters? Armed with massive troves of consumer, demographic, and voting data, Ghani and his team set out to investigate. However, they faced one crucial difference. In the supermarket project, all of the available data related precisely to the shopping domain. They studied shopping patterns to predict (and influence) what people would buy. But in politics there was very little relevant data available. Data teams for both campaigns needed proxies, and this required research.

They started out by interviewing several thousand people in great depth. These folks fell into different groups. Some cared about education or gay rights, others worried about Social Security or the impact of fracking on freshwater aquifers. Some supported the president unconditionally. Others sat on the fence. A good number liked him but didn't usually get around to voting. Some of them—and this was vital—were ready to contribute money to Obama's campaign.

Once Ghani's data team understood this small group of voters, their desires, their fears, and what it took to change their behavior, the next challenge was to find millions of other voters (and donors) who resembled them. This involved plowing through the consumer data and demographics of the voters they had interviewed and building mathematical profiles of them. Then it was just a matter of scouring national databases, finding people with similar profiles, and placing them into the same buckets.

The campaign could then target each group with advertisements, perhaps on Facebook or the media sites they visited, to see if they responded as expected. They carried out the same kind of A/B testing that Google uses to see which shade of blue garners more clicks on a button. Trying different approaches, they found, for example, that e-mail subject lines reading only "Hey!" bugged people but also led to more engagement and sometimes more donations. Through thousands of tests and tweaks, the campaign finally sized up its audience—including an all-important contingent of fifteen million swing voters.

Throughout this process, each campaign developed profiles of American voters. Each profile contained numerous scores, which not only gauged their value as a potential voter, volunteer, and donor but also reflected their stances on different issues. One voter might have a high score on environmental issues but a low one on national security or international trade. These political profiles are very similar to those that Internet companies, like Amazon and Netflix, use to manage their tens of millions of customers. Those companies' analytics engines churn out nearly constant cost/benefit analyses to maximize their revenue per customer.

Four years later, Hillary Clinton's campaign built upon the methodology established by Obama's team. It contracted a microtargeting start-up, the Groundwork, financed by Google chairman Eric Schmidt and run by Michael Slaby, the chief technology officer of Obama's 2012 campaign. The goal, according to a report in Quartz, was to build a data system that would create a political version of systems that companies like Salesforce.com develop to manage their millions of customers.

The appetite for fresh and relevant data, as you might imagine, is intense. And some of the methods used to gather it are unsavory, not to mention intrusive. In late 2015, the *Guardian* reported that a political data firm, Cambridge Analytica, had paid academics in the United Kingdom to amass Facebook profiles of US voters, with demographic details and records of each user's "likes." They used this information to develop psychographic analyses of more than forty million voters, ranking each on the scale of the "big five" personality traits: openness, conscientiousness, extroversion, agreeableness, and neuroticism. Groups working with the Ted Cruz presidential campaign then used these studies to develop television commercials targeted for different types of voters, placing them in programming they'd be most likely to watch. When the Republican Jewish Coalition was meeting at the Venetian in Las Vegas in May 2015, for instance, the Cruz campaign unleashed a series of web-based advertisements visible only inside the hotel complex that emphasized Cruz's devotion to Israel and its security.

I should mention here that not all of these targeting campaigns have proven to be effective. Some, no doubt, are selling little more than snake oil. The microtargeters, after all, are themselves marketing to campaigns and political action groups with millions of dollars to spend.

They sell them grand promises of priceless databases and pinpoint targeting, many of which are bound to be exaggerated. So in this sense the politicians not only purvey questionable promises but also consume them (at exorbitant expense). That said, as the Obama team demonstrated, some of these methods are fruitful. And so the industry —serious data scientists and hucksters alike—zeros in on voters.

Political microtargeters, however, face unique constraints, which make their work far more complex. The value of each voter, for example, rises or falls depending on the probability that his or her state will be in play. A swing voter in a swing state, like Florida, Ohio, or Nevada, is highly valuable. But if polls show the state tilting decisively to either blue or red, that voter's value plummets, and the marketing budget is quickly shifted toward other voters whose value is climbing.

In this sense, we can think of the voting public very much as we think of financial markets. With the flow of information, values rise and fall, as do investments. In these new political markets, each one of us represents a stock with its own fluctuating price. And each campaign must decide if and how to invest in us. If we merit the investment, then they decide not only what information to feed us but also how much and how to deliver it.

Similar calculations, on a macro scale, have been going on for decades, as campaigns plot their TV spending. As polling numbers change, they might cut ads in Pittsburgh and move those dollars to Tampa or Las Vegas. But with microtargeting, the focus shifts from the region to the individual. More important, that individual alone sees the customized version of the politician.

The campaigns use similar analysis to identify potential donors and to optimize each one. Here it gets complicated, because many of the donors themselves are carrying out their own calculations. They want the biggest bang for their buck. They know that if they immediately hand over the maximum contribution the campaign will view them as "fully tapped" and therefore irrelevant. But refusing to give any money will also render them irrelevant. So many give a drip-feed of money based on whether the messages they hear are ones they agree with. For them, managing a politician is like training a dog with treats. This training effect is all the more powerful for contributors to Super PACS, which do not limit political contributions.

The campaigns, of course, are well aware of this tactic. With microtargeting, they can send each of those donors the information most likely to pry more dollars from their bank accounts. And these messages will vary from one donor to the next.

■ ■ ■

These tactics aren't limited to campaigns. They infect our civic life, with lobbyists and interest groups now using these targeting methods to carry out their dirty work. In 2015, the Center for Medical Progress, an antiabortion group, posted videos featuring what they claimed was an aborted fetus at a Planned Parenthood clinic. The videos asserted that Planned Parenthood doctors were selling baby parts for research, and they spurred a wave of protest, and a Republican push to eliminate the organization's funding.

Research later showed that the video had been doctored: the so-called fetus was actually a photo of a stillborn baby born to a woman in rural Pennsylvania. And Planned Parenthood does not sell fetal tissue. The Center for Medical Progress admitted that the video contained misinformation. That weakened its appeal for a mass market. But with microtargeting, antiabortion activists could continue to build an audience for the video, despite the flawed premise, and use it to raise funds to fight Planned Parenthood.

While that campaign launched into public view, hundreds of others continue to hover below the surface, addressing individual voters. These quieter campaigns are equally deceptive and even less accountable. And they deliver ideological bombs that politicians will only hint at on the record. According to Zeynep Tufekci, a techno-sociologist and professor at the University of North Carolina, these groups pinpoint vulnerable voters and then target them with fear-mongering campaigns, scaring them about their children's safety or the rise of illegal immigration. At the same time, they can keep those ads from the eyes of voters likely to be turned off (or even disgusted) by such messaging.

Successful microtargeting, in part, explains why in 2015 more than 43 percent of Republicans, according to a survey, still believed the lie that President Obama is a Muslim. And 20 percent of Americans believed he was born outside the United States and, consequently, an illegitimate president. (Democrats may well spread their own disinformation in microtargeting, but nothing that has surfaced matches the scale of the anti-Obama campaigns.)

Even with the growth of microtargeting, political campaigns are still directing 75 percent of their media buy, on average, to television. You might think that this would have an equalizing effect, and it does. Television delivers the broader, and accountable, messaging, while microtargeting does its work in the shadows. But even television is moving toward personalized advertising. New advertising companies like Simulmedia, in New York, assemble TV viewers into behavioral buckets, so that advertisers can target audiences of like-minded people, whether hunters, pacifists, or buyers of tank-sized SUVs. As television and the rest of the media move toward profiling their viewers, the potential for political microtargeting grows.

As this happens, it will become harder to access the political messages our neighbors are seeing—and as a result, to understand why they believe what they do, often passionately. Even a nosy journalist will struggle to track down the messaging. It is not enough simply to visit the candidate's web page, because they, too, automatically profile and target each visitor, weighing everything from their zip codes to the links they click on the page, even the photos they appear to look at. It's also fruitless to create dozens of "fake" profiles, because the systems associate each real voter with deep accumulated knowledge, including purchasing records, addresses, phone numbers, voting records, and even social security numbers and Facebook profiles. To convince the system it's real, each fake would have to come with its own load of data. Fabricating one would require far too much work for a research project (and in the worst-case scenario it might get the investigator tangled up in fraud).

The result of these subterranean campaigns is a dangerous imbalance. The political marketers maintain deep dossiers on us, feed us a trickle of information, and measure how we respond to it. But we're kept in the dark about what our neighbors are being fed. This resembles a common tactic used by business negotiators. They deal with different parties separately so that none of them knows what the other is hearing. This asymmetry of information prevents the various parties from joining forces—which is precisely the point of a democratic government.

This growing science of microtargeting, with its profiles and predictions, fits all too neatly into our dark collection of WMDs. It is vast, opaque, and unaccountable. It provides cover to politicians, encouraging them to be many things to many people.

The scoring of individual voters also undermines democracy, making a minority of voters important and the rest little more than a supporting cast. Indeed, looking at the models used in presidential elections, we seem to inhabit a shrunken country. As I write this, the entire voting population that *matters* lives in a handful of counties in Florida, Ohio, Nevada, and a few other swing states. Within those counties is a small number of voters whose opinions weigh in the balance. I might point out here that while many of the WMDs we've been looking at, from predatory ads to policing models, deliver most of their punishment to the struggling classes, political microtargeting harms voters of every economic class. From Manhattan to San Francisco, rich and poor alike find themselves disenfranchised (though the truly affluent, of course, can more than compensate for this with campaign contributions).

In any case, the entire political system—the money, the attention, the fawning—turns to targeted voters like a flower following the sun. The rest of us are virtually ignored (except for fund-raising come-ons). The programs have already predicted our voting behavior, and any attempt to change it is not worth the investment.*

This creates a nefarious feedback loop. The disregarded voters are more likely to grow disenchanted. The winners know how to play the game. They get the inside story, while the vast majority of consumers receive only market-tested scraps.

Indeed, there is an added asymmetry. People who are expected to be voters but who, for one reason or another, skip an election find themselves lavished with attention the next time round. They still seem to brim with high voting potential. But those expected not to vote are largely ignored. The systems are searching for the cheapest votes to convert, with the highest return for each dollar spent. And nonvoters often look expensive. This dynamic prods a certain class of people to stay active and lets the rest lie fallow forever.

As is often the case with WMDs, the very same models that inflict damage could be used to humanity's benefit. Instead of targeting people in order to manipulate them, it could line them up for help. In a mayoral race, for example, a microtargeting campaign might tag certain voters for angry messages about unaffordable rents. But if the candidate knows these voters are angry about rent, how about using the same technology to identify the ones who will most benefit from affordable housing and then help them find it?

With political messaging, as with most WMDs, the heart of the problem is almost always the objective. Change that objective from leeching off people to helping them, and a WMD is disarmed—and can even become a force for good.