CHAPTER 17

On the Limits of Quantification

What You'll Learn

- Data and quantitative evidence don't tell us everything we need to know to make decisions.
- Sometimes the relevant evidence is inconclusive or non-existent, but this doesn't necessarily mean the right decision is to do nothing or to stick with the status quo.
- If we're not careful, quantification can have unintended ethical and equity implications.
- The data doesn't tell us what our goals are. Decisions must be made by thinking about both the effects of our actions and also our values.

Introduction

Data and quantitative analysis are full of promise for improving our lives and the world. But they have limits. We've seen many examples of the ways in which we must think clearly to use evidence well. If you mistake correlation for causation, ignore reversion to the mean or the over-comparing/under-reporting problem, try to establish correlation without variation, or pretend the data speak for themselves rather than viewing thinking and data as complements, quantification can lead you astray—resulting in worse, rather than better, decisions. Avoiding these pitfalls is exactly why we've worked so hard together to learn to think clearly about data and evidence.

We want to end by reflecting a bit on some slightly different limits of quantification and evidence-based decision making. These are not limits that come from a lack of clear thinking about some specific piece of quantitative analysis. Rather, they have to do with the realization that, important as evidence is, there is no such thing as a purely evidence-based decision. This is true for at least two reasons.

First, for many critical decisions, credible evidence is limited or even non-existent. But decisions still must be made. Indeed, even the decision not to act is a decision. So, it is important to think clearly about what we do when faced with an absence of evidence.

Second, the right decision can never be identified by evidence alone. Evidence is meant to be a tool used in service of our goals and values. But sometimes it seems like the tail wags the dog—our values become subservient to the dictates of quantification.

This is a dangerous mistake that must be guarded against through vigilance and clear thinking.

Decisions When Evidence Is Limited

There's an old parable that goes something like this. A drunk man is stumbling on the sidewalk looking for his keys under a lamppost. A passerby asks what he's doing, and the man responds, "Looking for my keys." The passerby inquires, "Where did you last see them?" to which the man answers, "I think I dropped them in the park across the street." The passerby reasonably asks the man why he is looking under the lamppost if he dropped his keys in the park across the street, and the man replies, "It's dark over there, I can't possibly find them in the dark! This is where the light is."

Clichéd though it may be, this parable illustrates an important point about quantification. We look where the light is. Not everything can be easily measured or quantified. And so, the analogue of looking where the light is in our data-driven world is narrowing our frame of vision to focus only on those things where quantitative evidence is available.

But such a narrowing poses real risks. First, we might end up simply ignoring crucially important problems because we don't see how to make an evidence-based decision. The fact that quantitative evidence isn't available to answer a question doesn't mean the question is unimportant or safely ignored. Second, the demand for evidence has the potential to create a kind of status quo bias. When someone says, "There's no evidence for that action," they might mean two different things. They could mean that lots of well-powered, well-designed studies have looked and turned up no evidence. But they could also mean that the action has never been studied (or even tried) before, so there is literally no evidence one way or the other. In the former case, it may be reasonable to not take the action. But in the latter case, there is simply no evidence available to guide us. If there are other good reasons to believe acting makes sense, it would be a mistake to let the absence of evidence force you into sticking with the status quo. Let's see some examples of these risks.

Cost-Benefit Analysis and Environmental Regulation

For the U.S. government's Office of Information and Regulatory Affairs (OIRA) within the Office of Management and Budget (OMB), quantitative evidence is the law of the land. A quantitative cost-benefit analysis is required for many new regulatory actions taken by executive agencies, and OIRA can essentially veto such regulations if they are unsatisfied by the evidence.

As we have discussed, not everything can be easily quantified. But, without quantitative evidence, OIRA approval is typically a non-starter. As a result, like the drunk man searching under the lamppost, regulators are forced to focus on those areas where quantification is possible, whether or not those areas are the places most in need of their attention.

Lisa Heinzerling, former head of policy at the Environmental Protection Agency (EPA), describes the bleak terms in which a former EPA staffer put it: "We're constantly asking ourselves not, 'Is this the right thing for environmental protection?' but, 'How can we make this acceptable to OMB?' "

In some sense, of course, requirements to quantify must frustrate regulators. The point of such requirements is to change the kind of regulation we get by changing regulators' behavior. The concern, however, is that these sorts of requirements don't simply prevent the EPA (and other agencies) from creating regulations for which the cure is worse than the disease. They also distort incentives in ways that narrow our field of vision. The mandate to quantify discourages agencies from bothering to work on regulations for which there are good arguments, but for which it is impossible, or too expensive or impractical, to quantify the costs and benefits. For example, in a typical EPA report about the regulation of an environmental contaminant, the regulators might discuss diseases and health conditions that they believe are affected by the contaminant. However, those diseases will only be incorporated into the cost-benefit analysis if we have a way to estimate the effect of the contaminant on disease risk and we have quantitative estimates of the monetary costs of the disease. And if they aren't included in the cost-benefit analysis, they won't have much sway over OIRA's decision making.

A well-known example is the controversy over the EPA's decision to tighten regulations on arsenic in water in the early 2000s. The EPA report making the case for the regulation lists a vast array of diseases that arsenic is believed to increase risk of. These include bladder, kidney, lung, liver, and prostate cancer, as well as a variety of other diseases with cardiovascular, pulmonary, immunological, neurological, and endocrine effects. However, the EPA notes that, because of lack of data, "the quantified benefits" included in their analysis concern only the effects of arsenic on "bladder and lung cancers." The rest of the health benefits of reduced arsenic exposure cannot be quantified. To their credit, the EPA used qualitative information to estimate these broader effects. But because of the demand for quantification, such estimates were easily dismissed in the ensuing controversy.

Floss Your Teeth and Wear a Mask

Two examples slightly closer to home illustrate why there are often good reasons to act, even absent quantitative evidence.

Floss your teeth

For years, Anthony flossed his teeth thoroughly every day because his dentist-spouse told him to do so, and because he believed her when she said that flossing is good for his health. But then, back in 2016, in the name of evidence-based decision making, the *New York Times* published an article entitled "Feeling Guilty about Not Flossing? Maybe There's No Need." It suggested that diligent flossers like Anthony could stop with the nightly flossing hassle.

The article in question cited a meta-analysis of twelve randomized experiments in which researchers compared the effects of brushing and flossing to just brushing. The article reported that the study "found only 'very unreliable' evidence that flossing might reduce plaque." So there you have it. An absence of evidence for flossing.[1]

So why hasn't Anthony stopped flossing? One reason, as discussed in chapter 6, is that failure to reject the null is not proof of the null—that is, absence of evidence is not (conclusive) evidence of absence. Even if we have no statistically significant evidence that flossing reduces plaque, this doesn't mean that flossing has no effect. What if the studies have low statistical power because of small sample sizes or large numbers of

[1] The meta-analysis actually does find statistically significant evidence that flossing reduces gingivitis, so the experimental evidence in favor of flossing is perhaps stronger than the article suggests.

noncompliers? Perhaps they wouldn't have detected effects even if flossing does reduce plaque.

Another reason is that researchers haven't studied all of the outcomes of interest. For instance, the authors of the meta-analysis point out that none of the experiments assess longer-term effects, nor do they study a variety of important dental outcomes like tooth decay, tartar, or gum separation.

But even these limitations of the quantitative studies aren't the whole story. As with many decisions, while we lack sufficient quantitative evidence to answer all of our questions about the effects of flossing, there are non-quantitative arguments that are important to consider. Dentists provide compelling biological and mechanistic accounts of why they believe flossing is beneficial. And so, despite what the contrarian data journalists might say, we're pretty comfortable with the decision to floss despite the fact that the quantitative evidence isn't conclusive. There are good reasons to believe that flossing is beneficial, even absent a slam-dunk empirical study.

Wear a mask

At the time of this writing, our society is having a similar debate about the effects of wearing masks in the midst of the COVID-19 global pandemic. Just as with flossing, there are few compelling, high-powered experiments demonstrating that masks and facial coverings reduce virus transmission. There are some observational studies that are subject to the concerns about confounders raised in chapter 9. Some of these studies focus only on selected samples of people who come into health care facilities with symptoms, which, as discussed in chapter 16, also creates problems. As with flossing, when researchers try to conduct a randomized experiment, many of the people assigned to treatment will fail to comply, making it more difficult to assess the effectiveness of wearing masks. And furthermore, we probably need a very large sample size to obtain a reasonably precise estimate of the effect of masks or flossing.

Given the lack of definitive evidence on masks, many people—including Donald Trump and Mike Pence, then president and vice president of the United States—decided to forgo the hassle. Such skeptics sometimes make arguments like "There is no evidence that wearing a mask matters." But as with flossing, there are good theoretical and biological reasons to think that masks are effective. We know that coronavirus and many other viruses are transmitted through respiratory particles, and we have good physical evidence that masks mitigate the flow of some of these particles. Studies also find that people who wear masks are less likely to touch their eyes, nose, and mouth, a second reason why masks likely mitigate transmission.

Of course, we aren't certain that we know the right answer, and we hope further studies will improve our understanding of the effects of wearing masks. But the lack of definitive, quantitative evidence in favor of one decision is not a compelling reason to make a different decision—especially if there is also no clear evidence in favor of that different decision.

Good decision makers use quantitative evidence, but they acknowledge that the quantitative evidence only tells them so much. They don't ignore certain considerations just because we lack good quantitative estimates for those factors. They use the best available theory and data to form their beliefs, and they make the best decision they can, given their goals, values, and those potentially imperfect and uncertain beliefs.

That last sentence pointed to another important thought about evidence-based decision making. No matter how good the data analysis, evidence alone cannot tell you how to act. For that, you need to also think about your goals and values. We end the book with some reflections on the ways in which quantification and those values interact.

Quantification and Values

Quantitative evidence should help us make better decisions that advance our goals and values. But if we aren't careful, matters can get turned around—our goals and values can be shaped by the mandate to quantify, rather than quantitative evidence serving our goals and values.

We are going to think about two ways this can happen. The first is that quantitative tools can sometimes sneak values into our decision making that we don't agree with, without our noticing. The second is that the desire to quantify can push us to embrace values that we might otherwise reject.

How Quantitative Tools Sneak in Values

One risk of quantification, especially in an age when machine learning and algorithmic decision making are increasingly prevalent, is that objectionable values will creep into decisions without our noticing. For instance, an algorithm may exhibit racial or gender bias, even if no data on race or gender were used to create the algorithm. This raises important questions about equity, fairness, and justice that deserve our attention.

Predictive machine learning algorithms get used for all sorts of tasks in the contemporary world. Job placement websites use such algorithms to match job seekers to employers. Banks use them to evaluate credit worthiness. Social media platforms use them to decide what content and advertisements to feed to users. And judges use them to inform criminal sentencing decisions.

How can these algorithms end up yielding ethically troubling results? Machine learning algorithms are, more or less, just fancy ways of using correlations to make predictions. An algorithm that is race- or gender-blind, in the sense of not having access to data on race or gender, could nonetheless end up making predictions that treat people with different racial or gender identities differently. This could happen, for example, if the algorithm has access to data on variables that are correlated with race or if some of the inputs of the algorithm are themselves subject to bias. We already saw an instance of this kind of problem in chapter 2, when we discussed how using the correlation between Yelp reviews and health code violations to target inspections would sneak in racial bias. But let's consider another example.

Algorithms and racial bias in health care

In the United States, large health care providers have special programs designed to coordinate the care of people with complex health needs. Such programs are expensive. So the providers only want to enroll people who are likely to have the greatest care needs. To try to predict who those patients are, they use machine learning algorithms.

There is a strong positive correlation between health care costs and health care needs because sicker patients tend to receive more and more expensive treatment. And health care costs are easier to measure accurately than health care needs. In this study, the algorithm was asked to predict health care costs. In order to do so, in addition to data on health care costs, it was fed data on patients' past insurance claims, medical diagnoses, and medications. Importantly, the algorithm specifically did not receive any information about race.

A simple way to think about how this might work is by analogy to regression. Suppose we had data on lots of patients' health care costs in year t and their insurance claims, diagnoses and procedures, and medications in year $t-1$. We could run the following regression:

$$\begin{aligned}\text{Costs}_t =&\beta_0 + \beta_1 \cdot \text{Insurance Claims}_{t-1} + \beta_2 \cdot \text{Diagnoses and Procedures}_{t-1} \\ &+ \beta_3 \cdot \text{Medications}_{t-1}\end{aligned}$$

Doing so would give us estimated OLS coefficients $\hat{\beta}_0$ through $\hat{\beta}_3$.

When a new patient, i, comes along, we can predict that patient's future health care costs using this algorithm. We feed that new patient's particular values for insurance claims, diagnoses and procedures, and medications into our regression equation to get

$$\begin{aligned}\text{Predicted Costs}_i =&\hat{\beta}_0 + \hat{\beta}_1 \cdot \text{Insurance Claims}_i + \hat{\beta}_2 \cdot \text{Diagnoses and Procedures}_i \\ &+ \hat{\beta}_3 \cdot \text{Medications}_i\end{aligned}$$

This is more or less what a predictive machine learning algorithm is doing, but the algorithm's goal is typically somewhat different than minimizing mean squared error, and it considers more complex functions of the variables than does a linear regression.

A 2019 paper in *Science* describes a health care provider using such predicted values to sort patients. Patients with a score above some high threshold were immediately enrolled in the special program. Patients with a score above a lower threshold were referred to a physician for further screening.

Even though the predictive algorithm was race-blind, it turned out to systematically under-estimate how sick Black patients were relative to White patients. This is illustrated in figure 17.1. Health care needs, as predicted by the algorithm, are on the horizontal axis. A measure of active chronic health conditions, called a *comorbidity score*, is on the vertical axis. This is meant to be a measure of true health care needs. As you can see, for any given level of predicted health care needs, Black patients turn out to be sicker than White patients, on average. Thus, Black patients were systematically less likely to be enrolled in the special program than White patients of similar health.

What might be going on such that this race-blind algorithm is nonetheless giving racially biased predictions? One possibility is an omitted variable. That is, perhaps even conditional on past insurance claims, diagnoses and procedures, and medications, Black patients tend to be sicker than White patients for reasons not observed in the data. This could result in the algorithm systematically under-estimating the health needs of Black patients and over-estimating the health needs of White patients.

In this case, however, it seems something slightly different is going on. The health care provider had the algorithm predict health care costs because costs are easily measured and are highly correlated with health care needs. But this decision proved problematic. A systematic fact about the U.S. health care system is that less money is spent on Black patients on average than is spent on similarly sick White patients.[2]

[2]The Readings and References section will point you to a review article documenting the many ways in which there is bias and discrimination against Black patients in U.S. health care.

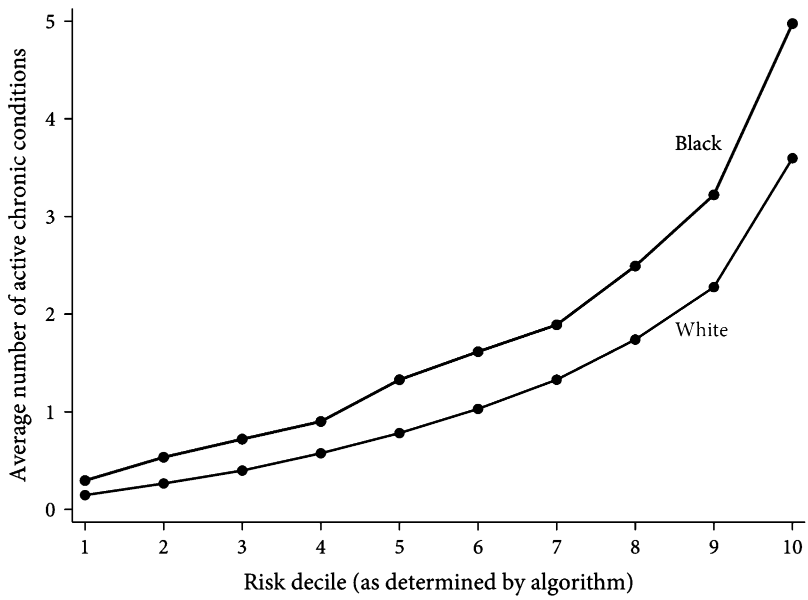

Figure 17.1. The relationship between the algorithmic prediction and actual health is different for Black and White patients.

This means using cost as a proxy for health care needs introduces racial bias into the supposedly race-blind algorithm. The algorithm correctly predicts that health care costs for a Black patient will be lower than for a White patient with similar characteristics (claims, diagnoses and procedures, medications). And that makes it look as though Black patients are healthier than equally sick White patients. Using the same inputs, but reformulating the algorithm to predict a measure of actual health rather than health-care costs, the authors of the *Science* article were able to eliminate the racial bias.

This example demonstrates how using quantitative tools can sneak values into our decision making that we might find objectionable. As the world becomes increasingly quantitative, it takes clear thinking and constant vigilance to make sure that our decisions are informed by data, but that the values driving those decisions are our own.

This brings us to our next topic, the ways in which quantification can shape the values that we think of as our own in potentially troubling ways.

How Quantification Shapes Our Values

Moral and political philosophers describe a rich variety of ethical concerns that one might consider when evaluating the rightness or wrongness of a decision. For instance, there are good and convincing arguments for various rights and duties, such as the right to control one's own body or the duty not to forcibly coerce one's fellow humans. A reasonable person might maintain that good policy or good actions must respect, or even promote, such rights and duties, even if violating them would lead to higher total material well-being in the world. This, for instance, is a position often held by principled opponents of the death penalty or torture or stem-cell research.

There are also good and convincing arguments for having concerns not only with total well-being but with the distribution of well-being. A reasonable person might, for instance, be prepared to accept lower total well-being in society in exchange for greater equality.

But most quantitative policy analysis is rooted in *welfarism*, the view that policies should be evaluated on the basis of their implications for human well-being. Moreover, one welfarist standard predominates over all others: *utilitarianism*—the view that policies should be evaluated on the basis of their implications for the sum total of human well-being, regardless of its distribution. And not just utilitarianism, but what we might call *crass utilitarianism*—one that defines well-being almost entirely in terms of material costs and benefits such as economic prosperity, health, and other factors that are (relatively) easy to quantify by assigning a monetary value.

An ethical position consistent with quantifying consequences is, in principle, quite flexible; it need not be crassly utilitarian. We can put a value on various non-material factors such as rights, duties, responsibility, dignity, or what have you. Moreover, once you know the quantitative effects of a policy on people's well-being, you can introduce all sorts of equity considerations into policy evaluation. We could, for example, after quantifying all the effects, define the best policy as the one that maximizes total well-being, subject to the constraint that everyone is above some minimal threshold.

What crass utilitarianism has going for it over all other normative frameworks—even other forms of welfarism—is that it lends itself easily to quantitative analysis. It is hard to figure out how to quantify the value of rights and duties or how to weigh equity considerations. It is much more straightforward—both conceptually and practically—to quantify material costs and benefits and then just add and subtract to figure out whether a policy or action is good or bad.

Indeed, crass utilitarianism is so easy to work with that it has become a part of the standard assumptions in the background of many quantitative analyses, especially in discussions of public policy. The process of trying to maximize net well-being—ignoring questions of rights, duties, responsibilities, equity, dignity, and so on—is often so ingrained in our practice and thought that we hardly even notice we are doing it. We simply take for granted that a good policy is one that maximizes benefits minus costs.

Notice what this means. The goals and values we pursue are shaped in a deep way by the dictates of quantification. We don't quantify because we are utilitarians. We are utilitarians because we quantify.

What's the problem with allowing a materialist utilitarianism to define our goals? In partial answer, we'd like to tell you a story.

Ethan once attended an academic presentation on the effects of removing children from abusive homes and putting them into foster care. The presenter found that children from abusive homes are, on average, better off in foster care. Moreover, the benefits for the children appear to exceed the costs of providing foster care. Therefore, the researcher concluded that we should remove kids from these abusive homes.

This seems like a great example of data leading to better policy decisions. We can quantify the benefits to children and choose policies that make them better off. We can even show that the benefits to the children exceed the dollar costs to society. So it looks like a clear win. Fantastic.

One of the attendees, someone who has held several senior positions in government, objected. The main critique was that the researcher had not estimated all of the relevant costs and benefits to make a policy recommendation. Specifically, what if the abusive parents derive benefits from keeping their children (and presumably, continuing to abuse them)? If the value to them of keeping the kids is big enough, then might that not reverse the conclusions of the cost-benefit analysis?

Now, you might think that a reasonable response to this line of questioning would be something like, "Well, if one were an unflinching utilitarian that would be right. But there are other values, and personally, I think we ought not give a damn about whether abusive parents want to keep their children. We should focus on what's best for the kids and for the rest of society." But that, in fact, was not the presenter's response. Instead, the presenter conceded the point, acknowledging that he really couldn't say whether or not taking kids out of abusive environments was good policy without knowing how it affected those kids' parents.

Or consider another example. In the early 1990s, the chief economist of the World Bank, Larry Summers—former president of Harvard, President Obama's chief economic advisor, and President Clinton's treasury secretary—circulated a memo written by his staff. It contained the following thought:

> Shouldn't the World Bank be encouraging MORE migration of the dirty industries to the LDCs [Less Developed Countries]?. . . The costs of health impairing pollution depends on the foregone earnings from increased morbidity and mortality. From this point of view a given amount of health impairing pollution should be done in the country with the lowest cost, which will be the country with the lowest wages. I think the economic logic behind dumping a load of toxic waste in the lowest wage country is impeccable and we should face up to that.

That toxic dumping in low-wage countries has "impeccable economic logic" is an interesting assertion. Here are three claims, each of which seems to us correct:

1. It is probably the case that the average willingness to pay for avoiding a little more toxic waste is higher in rich countries than in poor countries.
2. Hence, moving some toxic pollution from rich countries to poor countries will increase net material well-being in the world.
3. If these are the only costs and benefits (e.g., we don't count allowing rich countries not to take responsibility for their own actions as a direct cost) and we are utilitarians, then doing so is good policy.

To call that chain of arguments "economic logic" is troubling, for at least the last step has nothing to do with economics; it has to do with values. And the assumption in the first step that it makes sense to value outcomes based on willingness to pay also builds in potentially troubling normative priorities. We suspect the value of a marginal dollar is lower for richer people. So richer people have a higher willingness to pay than poorer people for the same change in well-being, simply because they value money differently. This means that if we evaluate costs and benefits based on people's willingness to pay, we are implicitly assuming the well-being of the rich is more important than the well-being of the poor.

Despite the ubiquity of these kinds of issues, at times, quantitative analysts seem to lose sight of the fact that making decisions by comparing such measures of material costs and benefits isn't value-free. Michael Greenstone, a prominent energy and environmental economist, makes the issue particularly clear in an argument he offers for using cost-benefit analysis for policy decision making:

> I think once we leave cost benefit analysis, then things start to bleed in often, not always, but often into moral decisions. And the deep problem, from my perspective, about moral-based decision making on a lot of these matters is that your morals aren't my morals, and a third person's morals are different than both of

> our morals. And then there's really no bounds on decision making...I have no confidence that what's good for all of society will be the result.

Greenstone is making an important point. By quantifying costs and benefits we put bounds on decision making. Such constraints can be very valuable. But he pushes the argument too far when he suggests that cost-benefit analysis takes subjective moral opinions out of the equation or that there is some objective scientific way to judge what's good for society without first making a set of value judgments that cannot be determined by evidence alone.

Of course, we can think of lots of reasons that, even if it passes a cost-benefit test, shifting toxic waste from rich to poor countries doesn't seem like good policy. Maybe we value fairness, justice, and economic mobility, such that we don't think dumping toxic waste on the poorest countries is a good idea. Maybe we think rich countries should take responsibility for their own actions. Maybe we don't want to live in the kind of world where rich people can simply pay for the right not to be affected by the pollution that is a byproduct of the economic activity from which they benefited. Maybe we think that the fact that poor people are less willing to pay to avoid toxic waste should be interpreted to imply not that their lives are less valuable than rich people's but that money is the wrong way to measure worth. Greenstone's point is right—reasonable people may disagree on all of these moral judgments. But reasonable people may also disagree on whether maximizing benefits net of costs justifies shipping waste from rich to poor countries. We must not let the fact that costs and benefits measured in terms of willingness to pay are relatively straightforward to quantify, while some other values are harder to quantify, fool us into thinking that evaluating things this way is objective science, while everything else is subjective value judgments. It all involves value judgments.

To be clear, we don't mean to suggest that there are no arguments in favor of the views expressed in Summers's memo. Suppose Summers is correct that transferring toxic waste from the rich to the poor will increase net material well-being. Then the rich might be able to more than compensate the poor for taking the toxic waste, leaving both parties better off as a result of a trade. Hence, if we have the technological ability and political will to get the poor to take the toxic waste and to get the rich to pay them for doing so, we might be able to create a win-win situation.

But of course, there are lots of other moral arguments to consider as well. In our view, it should matter how this decision is made. We would think very differently about poor countries agreeing to accept toxic waste in exchange for compensation than we would about an economist in a rich country making the decision and telling the poor country that they're better off. But Summers's memo doesn't even express concern about whether the poor countries will be compensated or agree to this arrangement. The crass utilitarian argument, on its own, appears to suffice. To his credit, in later discussions of the toxic waste memo, Summers expressed a different view. For instance, in a 1998 interview with the *New Yorker* he said, "The basic sentiment that it is good to ship toxic wastes to poor countries is obviously all wrong. Are there real issues about trade-offs between growth and the environment? Sure. But the way the thoughts were expressed wasn't constructive in any sense."

The cases of abused children and toxic dumping are interesting for several reasons. Quantification often pushes us toward crass utilitarianism, which can lead to ruthless and absurd conclusions. But the discipline of quantification also really does teach us something. For many people, the rich dumping their toxic waste onto the poor might not appear to be a policy idea with anything to recommend it. The exercise of quantifying and comparing costs and benefits forces us to see that there is a serious argument to be made for this policy (at least the version of it that involves consent and compensation), even if ultimately some of us come down on the other side.

In both cases, we believe we can (and should) realize some of the benefits of quantification—precision, weighing trade-offs, contestability. But both cases also illustrate a key concern. Rights, dignity, and fairness are hard to quantify. Material costs and benefits are easier. And so, in practice, the desire for quantitative evidence pushes us toward a focus on the kind of highly objectionable, crass, materialist utilitarianism that characterizes these stories. If we are to use quantitative analysis for good, we must strive toward a practice in which data and quantitative tools help us estimate important quantities without distorting the goals and values against which we evaluate our choices.

Think Clearly and Help Others Do So Too

We'd like to conclude by urging you to use the tools and skills you've learned for good. Much of this book was about how thinking clearly can help you spot when someone is intentionally or accidentally misleading you with data. But a cynical reader could turn that noble mission on its head, using these tools as a recipe for misleading others who haven't learned to think as clearly. Unless sales of this book really go through the roof, most people you interact with won't notice if you assert correlation without variation, claim evidence for a causal relationship from a correlation that you know is confounded, or keep making comparisons until you find the conclusion you want and then only report that one. Please don't do that! Think about the larger quest for truth, and take your newfound responsibility as a savvy quantitative thinker seriously. Be transparent about the strengths and weaknesses of the evidence you bring to bear, whether that is evidence you created through your own analysis or read about in someone else's. In so doing, you can help others, as well as yourself, think clearly with data.

But, most importantly, take a moment to appreciate how hard you have worked and how far we have come together. You are now a member of a small but growing group of people in the world who can think clearly about the problem of selecting on the dependent variable, the difference between statistical and substantive significance, reversion to the mean, publication bias, the sources of cosmic habituation, the relationship between correlation and causation, foundational ideas about research design, and so much more. These are fundamental conceptual understandings that will serve you well forever, even if you never run another regression. Because we all now live in a time in which thinking clearly with and about data is absolutely essential for anyone who wants to understand the world and make it a better place.

Exercises

17.1 Your friend Andy has noticed that every time he eats pancakes for breakfast, he does well on his exams. Therefore, he has decided that his diet will, from here on out, consist entirely of pancakes. On the basis of the lessons from the entire book, list at least four things wrong with your friend's reasoning.

17.2 You are the mayor of a major city, and your staff presents you with a plan to provide more amenities to low-income neighborhoods. They tell you that their plan will cost $100 million, but they estimate that it will provide $200 million of economic benefits, so it's a no-brainer. What questions would you want to ask your staff before deciding to proceed with the plan?

17.3 Think of a decision in your life that you've made largely without the help of quantitative evidence (such as Anthony deciding to floss despite there being few compelling, quantitative studies on the topic). What factors led you to make the decision you did? Can you propose a quantitative study that would provide more compelling evidence? What would the evidence have to look like in order for you to change your decision?

Readings and References

You can read the EPA's report on the arsenic regulation, including the long list of non-quantifiable health impacts, in the federal register at https://www.govinfo.gov/content/pkg/FR-2001-01-22/pdf/01-1668.pdf.

The meta-analysis on the effects of flossing is

Dario Sambunjak, Jason W. Nickerson, Tina Poklepovic, Trevor M. Johnson, Pauline Imai, Peter Tugwell, and Helen V. Worthington. 2011. "Flossing for the Management of Periodontal Diseases and Dental Caries in Adults." Cochrane Database of Systemic Reviews, Issue 12. doi.org/10.1002/14651858.CD008829.pub2.

This research team updated the meta-analysis, adding three more experiments on flossing in 2019. See

Helen V. Worthington, Laura MacDonald, Tina Poklepovic Pericic, Dario Sambunjak, Trevor M. Johnson, Pauline Imai, and Janet E. Clarkson. 2019. "Home Use of Interdental Cleaning Devices, in Addition to Toothbrushing, for Preventing and Controlling Periodontal Diseases and Dental Caries." Cochrane Database of Systemic Reviews, Issue 4. doi.org/10.1002/14651858.CD012018.pub2.

For some empirical evidence on mask wearing and the spread of particulates, see

Sima Asadi, Christopher D. Cappa, Santiago Barreda, Anthony S. Wexler, Nicole M. Bouvier, and William D. Ristenparth. 2020. "Efficacy of Masks and Face Coverings in Controlling Outward Aerosol Particle Emission from Expiratory Activities." *Scientific Reports* 10, Article 15665. doi.org/10.1038/s41598-020-72798-7.

For evidence on the relationship between wearing a mask and touching your face, see

Yong-Jian Chen, Gang Qin, Jie Chen, Jian-Liang Xu, Ding-Yun Feng, Xiang-Yuan Wu, and Xing Li. 2020. "Comparison of Face-Touching Behaviors Before and During the Coronavirus Disease 2019 Pandemic." *JAMA Network Open* 3(7).

Tiffany L. Lucas, Rachel Mustain, and Robert E. Goldsby. "Frequency of Face Touching with and without a Mask in Pediatric Hematology/Oncology Health Care Professionals." *Pediatric Blood & Cancer* 67(9).

For a discussion of how race relates to disparate health outcomes in the United States, see

David Cutler, Adriana Lleras-Muney, and Tom Vogl. 2011. "Socioeconomic Status and Health: Dimensions and Mechanisms." *The Oxford Handbook of Health Economics*, Sherry Glied and Peter C. Smith, eds. Oxford University Press.

We referenced this study in our discussion of how quantitative tools can sneak in values:

Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Science* 366(6464):447–53.

You can find the quotation from Summers and a discussion of the toxic waste memo here:

John Cassid. 1998. "The Triumphalist." *The New Yorker*. July 6.

The quotation from Michael Greenstone is from this podcast:

"The Value of a Life" (episode 1). *Pandemic Economics*. Becker Friedman Institute. April 23, 2020. bfi.uchicago.edu/podcast/pandemic-economics-01.

The discussion of how quantification can shape our moral values draws heavily on

Ethan Bueno de Mesquita. 2019. "The Perils of Quantification." *The Boston Review*. March 11. https://bostonreview.net/forum/economics-after-neoliberalism/ethan-bueno-de-mesquita-perils-quantification.