Patrick Ball: I would say it's so much more bureaucratic than most people imagine. I don't know that it's all that hard, but it's very bureaucratic, and I think that's the piece that people generally misunderstand. People look at truth commissions and they think there's these big ideological conflicts going on inside the commissions. That's not always false, but those are rarely important in the day-to-day work. What's generally important are giant fights over the execution of the bureaucratic mandate, because it's very resource-intensive, and it's very difficult, and so people spend a lot of time fighting over what it is actually we're supposed to do as bureaucrats.

Particularly in South Africa, that was the major challenge, and it's confused people ever since. Every analyst that I read, pretty much, especially Americans, who've analyzed the Truth and Reconciliation Commission, misunderstands this. They really want to read the Commission's work as an extension of apartheid or as an effort by the African National Congress to establish their legitimacy as the new government, or create some sense of national consensus around the transition.

But internally, in the commission, it was just all about figuring out, how do we process these perpetrators in the amnesty queue in some kind of reasonably coherent way, is that judicial, is it non-judicial, what is the framework, what's the bureaucracy here? Not even the ideology, just what's the bureaucracy? What are the steps, who has to review what in order to make a finding? And then how do we listen to the victims in a systematic way? How do we tell their stories fairly? How do we go about getting victims appropriate compensation? How do we provide that, all the restitution stuff, and how do we do that in a reasonably just way?

While there may be occasional connections between the outcomes of these administrative or bureaucratic fights and larger political questions, no one is thinking about them internally, and I think the connections tend to be arbitrary. They don't go all in one direction or another.

This work was very hard, and I think we failed. Badly. I've worked with nine commissions, and the South African experience was more problematic than the others I've worked with: it failed to create an unequivocal truth, it failed to foster reconciliation, it failed to come up with a good restitution policy. Let me be clear: I was part of the failure, I made giant mistakes.

Q: TK

PB: Oh, because there were really nice hotels for journalists to stay in while they were there, and plus, everybody loves Desmond Tutu. That's why.

Q: TK

PB: I mean, especially everyone who comes from a kind of Christian perspective, which I really don't. So, if you were Christian, stuff he said about reconciliation and forgiveness made sense, but if you were Jewish, Muslim, or Hindu, or an atheist, which many South

Africans are, the Christian religious language was a distraction or incoherent or offensive or all of those.

Q: TK

PB: But some elites in the Afrikaner community, in the African community, and in the English-speaking white community were happy to sign on to this Christian understanding of reconciliation. 5 of the 17 commissioners were practicing religious figures. Christian religious figures. And so, the atheists, the Jews, the Hindus, and the Muslims kind of got shut out by the Christian talk. The overbearing Christianity started with Archbishop Tutu.

And Vice-Chair Alex Boraine, too. Boraine was less obviously religious, although he was an ordained practicing Methodist minister.

Q: TK

PB: Yeah, I agree. But they're appealing in a way that's bound to Western-ness. Tutu would say radical things. My favorite of his quotes is, when the white people came, we had the land and they had the bibles, now we have the bibles and they have the land. And there's a little bit of radical implication there—I mean, should we swap back?

Anyway, South Africa was a weak commission. It was weaker than the Liberian TRC, which faced a lot of serious challenges, especially with an inadequate budget. The Liberian TRC was similarly governed by a deeply Christian vision of forgiveness and reconciliation. But the Liberians benefited from the experience of many prior commissions processing thousands of narrative statements and so were able to build their data foundations on a more technically sound basis.

Q: TK

PB: Sure. My name's Patrick Ball. I am the director of research, Human Rights Data Analysis Group. Since 1991, I have built databases and conducted statistical analysis for variety of human rights projects, including 9 truth commissions. Various criminal tribunals, I've been an expert witness in a number of tribunals. I've been an advisor to a number of human rights missions and dozens and dozens of non-governmental human rights groups in over 30 countries.

Q: TK

PB: No, I was in El Salvador in 1991 for completely different reasons, and my partners there, a Salvadoran human rights group, asked if I knew anything about computers. I said, oddly enough, I used to make my living as a programmer. And they were like, "Oh, great. So, let's put together a project to map our 9,500 legal depositions that we've taken since 1977. Let's code those out." We were hand-classifying the stories, one by one, into tables. This is long before any natural-language processing was a glimmer in any of our eyes. These were DOS-based monochrome computers with 640 kilobytes of RAM. I know, this is Grandpa's computing talk, but we had only 20-megabyte hard disks.

I had to write SQL code on those, which actually worked even with those primitive computers. People at the NGO coded all those 9,500 hundred stories into tables, and we built a big database which we combined with a second database of the career histories of the 400 most senior officers in the Salvadoran military. And the idea was to link those, just by the unit having committed the violation and the date of the violation, to that segment of that guy's career, each of the segments of the careers of the officers in that unit at that time, and thereby create a dossier of violations for each officer. So, you're sort of inverting the way we understand the violations as connected to the officer, rather than to the victim in some sense. Although it's really both.

I think that El Salvador was probably the first really big use of data in human rights with this application. And by big, I just mean really almost any systematic use of data as data, rather than computers as fancy typewriters and filing cabinets, places to put text and findings, but rather structured records that you're moving around in a structured way with quantitative implications, well, I think the projects in El Salvador were probably there first.

Q: TK

PB: These folks in L.A. had thought it up. We didn't actually come up with the idea. We copied them, but we improved on their work, I think. So, this NGO in L.A., called El Rescate, had done an initial version of it. We copied their logic, but we built our own software. The El Rescate project was published, and then our project, and then the truth commission was the first modern truth commission, in 1993. And by modern, I mean the first one to take every interview they could possibly get their hands on and then render it as data. And they were the first one to try to understand a really massive number of testimonies.

Q: TK

PB: That's actually much later, I think. That's actually not until the aughts.

Q: TK

PB: I don't think it was one of the earliest truth commissions. The Argentines and the Chileans had come much earlier. But they were really just big law firms in some sense. But they were still fundamentally legal exercises, where they're investigating cases that should've been investigated by the judiciary. So, in Latin America, the legal system is such that cases are investigated by a judge, not by a prosecutor. And so, a complaint comes before a judge, either because police bring or individual citizens bring it, and then the judge is supposed to investigate. And judges have resources for this.

The Chilean and the Argentine commissions basically said here are a bunch of cases which should have been investigated. Although they counted some stuff up, but they didn't really think about the statistics or the implications of counting.

And then in El Salvador, what was really different is they talked to thousands of people, which is kind of remarkable, given that El Salvador is a very small country. Nobody really understood this at the time, that what we were doing was really innovative. And then they ended up with all this data, like, now what do we do? This guy, who later became a good friend of mine, Ken Ward, came and studied the database that I had written at the CDHES. He reworked the ideas we had figured out at the CDHES to use at the commission. And they coded out all the testimonies they'd taken, and they did some rudimentary statistical work on describing the data that they collected. It wasn't a bad database, pretty damn good, honestly, for 1991-1992.

And then a series of commissions that followed, Haiti and then Guatemala and South Africa in the '90s, followed similar logics, where we're looking at everything we can, and we're putting it together and doing descriptive understanding of what it is. And they all had changed the logic a little bit over time, from this notion of we've got a bunch of legal cases. The Salvadorans kind of refined that into, these are illustrative cases. These are cases which illustrate what it is we're trying to do in the search for the truth. The story was fundamentally about these few cases, though it was changing to be a bigger, more inclusive story, relative to what the Chileans and Argentines had done.

In Guatemala, it became much more ambitious, and we said, well, we're going to try to actually tell the whole story, whereas the South Africans were retreated a little bit, and kind of mixed the Chilean and Salvadoran models, with a timeline-based history, in which they extracted tiny little cases and did little bits of investigation on. But the investigations weren't very deep, and the selection of the periodization and the categorization, that didn't really make sense of South African history as well as it could have. Argument was spread thinly over time, and the statistical findings were not connected to the narrative history.

My main critique is that the South African commission enormously privileged white victims. In fact, it was a position that the archbishop insisted on, that all the public hearings should have a victim from each of South Africa's ethnic categories, which makes no sense, because those proportions of suffering were completely different: African people suffered a lot more than white people. But this was all about creating a sense that all South Africans suffered from apartheid (a position which Tutu argued explicitly in the opening pages of commission's final report). It was a kind of ideological statement, but it was a Christian statement intended to create a sense of shared suffering, rather than a nationalist vs. ANC or left-right kind of ideological statement.

Q: TK

PB: Right. So, anyway, that's the kind of thing. And then we hit the wall. In 1999, we realized that these databases were, in some sense, making less truth than we started out with—basically because we thought the raw data would tell the story. And the reason is that we don't know what we don't know: the raw data tells a very specific part of the story, but you don't know what part it is. And this is the fundamental problem faced by statistics in the 20th century: how do you understand the world based on only a small piece of data? How can you account for what is unobserved?

Since 1999, my work has become, more and more obsessively, about what we do not observe. How do we do work with the data we don't have? How do we model for it and understand patterns and relationships quantitatively in ways that take into account what we do not observe? And there are a bunch of models to do this, and we use these different models in different places, and Timor and Peru were the most successful, I think. In Guatemala, we did the first version of this thinking.

Q: TK

PB: And that is the same project as we did earlier in Guatemala, but the application there was to the trial, which came in 2013. The trial of the former dictator Ríos Montt. So yeah, that's good, you got that piece, but that logic, the problem of trying to figure out what you don't observe is, I think, the much bigger human rights data question, and it has been the question that I've been interested in for 20 years now.

Q: TK

PB: So, you go to highland Peru, and you go and you ask a bunch of people, hey, who died in the war? And people are like, oh, Jose and Jorge died. Okay, and you write that down... Jose and Jorge, thanks. Who killed them? Oh yeah, Sendero killed Jorge, and the army killed Jose. Yeah, okay. So you go home, and then you have to ask yourself, did I get all the names of the dead, or were there some of the families who decided to ignore me? Put yourself in the shoes of people in the area you were surveying and imagine their thinking: I don't know those people asking all these intrusive questions. Why am I going to talk to them? What good might come of it, versus what bad things might happen? Let's just walk away, we're busy, we have other things to do, we're not spending any time talking to those people. But nonetheless, there were members of their family who were killed. So, how do we know how many of those there are, relative to the ones we have? That's one of the problems: underreporting that leads to selection bias.

The other problem is, okay, well, we've documented Jorge and Jose, one killed by Sendero, one killed by the state. Does that mean that the proportion of responsibility is 50/50? That's what we observe, because we got one victim from one side, and one victim from the other side. Does that mean Sendero and the state committed equal numbers of homicides? Or is it the case that if we kept asking people, it would turn out that most of the people that we didn't hear about in the earlier projects were really killed by Sendero?

We learned that we heard about most of the victims of the state, but we heard about a much smaller fraction of the people killed by Sendero. In the data collection prior to the commission's work, people thought, if the state killed us, we might get some kind of accountability from state, so there's value in denouncing the killings by the state. But people knew that during the civil war, the state can't do anything if Sendero killed my brother. The state is not going to take responsibility for that, of course, because it wasn't their fault. So, if they're not going to give me justice for my brother getting killed by Sendero, so why should I bother? This is an entirely reasonable calculation, because there's risk involved in disclosing in this stuff. Some of the perpetrators are still around

and might retaliate against people talking about their crimes. And there's retraumatization. Nobody likes to talk about this stuff.

And so, you get these weird distortions in your understanding about these key questions when you're talking about the truth. In Peru, perhaps the most important statistical question about the past is what are the proportions of responsibility between the state and Sendero? You can get that totally wrong, especially if you freeze the picture in 2003, where the only project to really document Sendero killings was the truth commission. As I mentioned earlier, all the prior projects had pretty much only focused on the state because they were projects trying to hold the state accountable. It was not an accident that that's what they focused on. They didn't purport to be covering a statistically representative sample of the world, but people nonetheless read them that way and misunderstood reality. And this is often this case.

What data collection does may be perfectly reasonable, but how people interpret that data collection may be naïve. It's very common.

Q: TK

PB: Yeah, I've never really been very fond of that distinction.

Q: TK

PB: Yeah, I would be pretty strict on the statistical side. Statistical significance is a term of art. If I were examining a problem in an epidemiological assessment like that, I would probably not choose that particular statistical framing to study it. Women who live in places where there is no contamination have miscarriages and children with heart defects. The question is, whether women who lived in an area where there's contamination have more or fewer negative birth outcomes relative to women who live in uncontaminated areas? How can we determine that? And if the rate's the same, it's hard for me to attribute those issues to the contamination. The contamination may be a problem in other ways, but linking it to these specific health outcomes...I am committed to science and the way that science understands things, and epidemiology has ways to understand questions like that.

Now, whether or not it works in court is a different problem. I have been in court, and court is not a great truth-seeking mechanism. With some judges and prosecutors it can work, and we've had some really good legal outcomes where judges dug into the scientific evidence, but on the whole, legal procedure is designed for justice, not for scientific truth. Court is not usually a good place for science.

That said, science is a good place for science. So, if the science is clear, I'd be happy to make the attribution.

Q: TK

PB: Of course. Scientists never accept things blindly unless they're terrible scientists. You don't do that. You build all kinds of elaborate tests. The good part of science, for me, is

not always the discovery part. Instead it’s trying to disconfirm a finding, and then failing to disconfirm an idea about how something works. This is always exciting, because you created a test that your analysis might fail, and if your finding fails to fail the test, you're like, okay, the double negative means that I've got one more tiny crumb of assurance now. But it's a crumb, and you got to build a lot of crumbs, so you build the other tests, and you try to figure out, how many different ways can you weaken the assumptions or reverse the assumptions and see if the results come out the same?

Q: TK

PB: People often look at data and say, well, that data's wrong, and therefore, we can't learn anything from data. And I say well, sure, the data is wrong. Data is nearly always wrong. But that doesn't mean you can't learn something from it. You can still learn things from wrong data, but you have to start by figuring out how it’s wrong. That’s a mathematical process, not hand-waving speculation about how one or another project did better or worse fieldwork. And if you have lots of wrong data, you may be able to make estimates, you may be able to build a model with that data that corrects for the selection bias, and if all the assumptions are right, you may have a finding that is no longer wrong. It might even be right. This is hard.

The statistical stories we want to tell about the world with data are generally comparative stories. So, we want to say, well, homicides were more frequent in April than they were in March. Killings went up. So, we've got a comparison between two categories, time categories, March, April. We put everything in the March box that happened in March, put things in the April box that happened in April, and we then look. Which is greater? Say, well, it went up, okay? Well, did it go up because you had more people working for you in April, so you collected more data? You don't actually know if there was more in March or April, you just know you got more data in April. The statistical question you have to ask is what's the fraction of the universe that you are able to observe in each period?

People who treat data as if it were true make the naïve assumption that the fraction of the universe they observe, across their various categories that they're comparing, is constant. That's unlikely, really unlikely. One of my mantras is that data's always lying to you. And it's lying to you in this very specific way. It's lying to you because this probability of observation or the fraction of the reality that you observed with your data-capture mechanism varies wildly.

The phrase that my colleague Megan Price and I use for this is the documentation dynamics. The dynamics of documentation vary in ways that are different from the conflict dynamics. So, as violence is getting worse in April, it might be that, simultaneously, your ability to observe is degrading because a bunch of good people working for you got killed, or they (quite reasonably) fled the worsening violence. And so, it looks like violence is going down when it's really going up, specifically because the conflict dynamics are influencing the observed documentation dynamics in a way that lead you to the wrong answer.

One of the most important assumptions that I try to unpack for people is this assumption of constant observability. So, what that means is that you cannot just believe the data. You have to do something with the data before you start making conclusions or arguments from it. You have to create a model, and the model has to be specifically about the documentation dynamics, or the sampling structure, the sampling probabilities. What do you do? Well, there are a bunch of models you can use, and I don't want to go down that path too far; we have nine minutes before I have to go.

But the next assumption that I think people make that's often incorrect is that, after we've made an estimate, we've done these corrections, we've modeled for documentation dynamics, we made estimates, people still think, that because we have biased data, our results are biased. But what we've just done is model for that bias. We've actually done everything we know how to do with tools and science to correct for those biases that you're worried about. We still might not have fixed it! But if the assumptions are right and the model fits—big ifs—we're reasonably likely to be right.

People tend to have excessive confidence in their data, and excessive skepticism about models. I think that should be reversed. Good models make explicit what we have to assume to know the truth. Once people get their heads around, oh, the data is being produced in some human way, which responds to human strengths and weaknesses. It's not produced in some transparent way by the reality that we think we're monitoring. Rather, it's an effect of our logistics, our funding, the trust we have from communities we're talking to, our informational capacity to manage the information, all of those things govern what we're able to see, and it varies all over the place. So, different parts of an organization, different parts of the country, at different times, at different capacities. And we do a better or worse job based on those capacities.

When I worked for the U.N. in Congo, every once in a while there would be some sort of scandal about some of the peacekeeping troops, or one of the troops, having a sexual relationship with some local person. And the allegation would be either that the relationship was coercive or that the local person was underaged or some other serious problem. So, you have one of these scandals, and then the local people won't talk to you for a few months, and as a consequence, your reported homicides from that region just fall to zero. Well, does that mean that homicide fell to zero? No. Maybe they went up, maybe they went down, but they almost certainly didn't stop entirely. Maybe the militias showed up, because now no one wants to talk to the UN, so the militia showed up, and homicides went up. But due to the collapse in trust from the community, you don't get any data.

Q: TK

PB: A little, but some of it's coincident with power dynamics, but the power dynamics might not be the obvious ones. So, in my experience, most of the power dynamics that really affect the way these kinds of institutions work are internal to the organization. So, these bureaucratic bun-fights inside the organizations that deeply influence what the organization chooses to do. So, I've given you a couple of examples. One is the fights, bureaucratically, within the South African TRC, about what it meant to certify someone

as a victim so they get compensation. These were not trying to privilege one community over another, these were just different, and sometimes, unfortunately, incompatible ideas about how the administrative process should work in order to be effective and fair. This kind of conflict happens even when, or especially when, everyone involved is of good faith and trying to do the right thing. Even though I disagreed with some decisions that some commissioners made, I recognized that they were trying to do the right thing.

I think there are always power dynamics inside these kinds of arguments, but there may not be left/right, global north/global south elites within the country. They may be very internal to these systems being built, and you may really have to get inside to figure out what's the bureaucratic fight. Sometimes people will link their conflict to external politics because they legitimize their position, and they bring power and attention to them. No matter what people say, that doesn't necessarily mean that the internal fights have much to do with the external politics, and it can be very confusing.

Q: TK

PB: One of the things that I would love to hear is if you could kind of express is that the best commissions, really the best commissions, have an intellectual richness I've never experienced in any other project. The Peruvian and Guatemalan commissions are the ones that really stand out to me. There's a sense of really the best social science that I've ever heard, seen of, or even imagined getting done in these contexts. Deep, people-centered histories, the best combination of local narrative, big-picture structural explanations, and quantitative reasoning, mixed together in this amazing, fantastic history. I will remember for the rest of my life some of the moments in the Peruvian and Guatemalan commissions where I felt like the people in the room were all so deep in these different areas, but none of us had disciplinary fences up. We weren't quibbling over qualitative versus quantitative, juridical definitions or local meaning, or social history versus structural history, none of that was relevant. All we wanted to do was figure out what the story was and tell it, using all the disciplines we had.

And it was fantastic. They've been the most intellectually satisfying experiences of my career, and I've never found a context like that except in truth commissions. I mean, sometimes I've heard it a little bit in really good war crimes prosecutions, but there it's rare because it all has to be reframed in these legal and juridical mechanisms to get it in front of the judge, whereas in truth commissions, you're just telling a story. So, you have this incredibly exciting exchanges, where all the silly academic boundaries are discarded, and you're just trying to figure out the story.

So, I hope that you're able to express a little bit of that, just to say there is this possibility. It doesn't always make it to the report, but when it does, it's amazing. You have these moments of just incredible synergy and excitement and insight.

Q: TK

PB: I think I'm a little more skeptical about the truth commission model than I was five or ten or twenty years ago.

Q: TK

PB: Yeah, and doing it in a way that creates political consensus, so that you have political consensus at least internally about what you're trying to accomplish. You've invited to the table people with sufficiently diverse views such that at least most of the society will believe at least one of the commissioners. And then you tell a consensus story that everyone at the table's willing to sign off on.

Q: TK

PB: Narrowing the range of permissible lies?

Q: TK

PB: I don't actually think that phrase tells the story correctly. I loved it for a long time, that phrase, but I don't actually think it's what commissions do. I actually think that's what tribunals do. That's not what commissions do. What commissions do is they provide a kernel of a story that most people have consensus about. And so, it is converse of Ignatieff's contention, in some sense the affirmative version of his claim. I think it is trials that actually rebut the egregious lies. Trials don't usually get the whole truth, and they may well fail to engage or understand important evidence, but the specific facts on which judicial findings are made tend to be good facts.