NEW ZEALAND INFRASTRUCTURE COMMISSION
Te Waihanga

Episode five

TRANSCRIPT

Infrastructure for a better future

Benchmarking the Cost of Infrastructure

Peter Nunns
Director - Economics, Te Waihanga

Andreas Leed
Head of Data Science, Oxford Global Projects

Dirk Pöker
Research Analyst, Oxford Global Projects

Please note: the transcript has been edited to make reading as easy as possible.

Introduction: Welcome to the Te Waihanga 'Infrastructure for a Better Future' podcast. A series where we talk to experts both from here and overseas about the infrastructure challenges we are facing.

Peter Nunns: To solve New Zealand's infrastructure challenges we need to deliver quality infrastructure at an affordable price. So, how much does it cost to build infrastructure in New Zealand? And how do we compare with other high-income countries? That's the subject of a report that we published last month, entitled 'The lay of the land: Benchmarking New Zealand's infrastructure delivery costs'. In our report, we draw on research from Oxford Global Projects, who provide expert advice on infrastructure project management and benchmarking around the world. Both our report and the Oxford Global Projects companion report are available on our website, www.**tewaihanga**.govt.nz.

For this podcast we're joined by Andreas Leed and Dirk Pöker from Oxford Global Projects, who will discuss their findings and share some of their broader insights. Andreas is head of data science at Oxford Global Projects and Dirk is a research analyst and PhD candidate at the University of Alicante in Spain. Now, the analysis you did compares costs for seven types of infrastructure projects, urban and rural motorways, road tunnels, rail stations, electricity transmission lines, wind farms and hospitals around the world. What were some of your key takeaways from that analysis?

Dirk Pöker: Yes, our findings were actually quite intriguing. I think New Zealand as a nation is taking on the challenge of improving the efficiency of its infrastructure sector. But actually on analysis, it turns out that New Zealand does not have significantly high infrastructure construction costs across the board. Our benchmarking of international costs shows that it is only really when it comes to complex large-scale projects such as large motorways, with a lot of bridges, road tunnels, underground rail projects, that the costs do tend to be higher than other high-income countries. We suppose that New Zealand's difficult terrain plays a significant

role as a cost driver. The cost to deliver smaller, standardised infrastructure on the other hand, like surface rail stations, electricity transmission lines, onshore wind farms, hospitals in New Zealand is similar when compared to other OECD countries.

Peter Nunns: That was the most the most interesting part of part of the work for me, right is seeing that differentiation between different types of projects. It's not all the same story. What do you think that means for the infrastructure sector?

Andreas Leed: I think I can chip in here and say that I think it is quite an interesting challenge for the infrastructure sector. To actually improve upon, especially the more complex, large-scale types of infrastructure, the government probably needs to drive the transformation. I know that reading your report, based on our report, you make the recommendation that the government should act as more of a sophisticated client when it comes to infrastructure projects. I think we absolutely agree with this point. A couple of ideas for acting as a sophisticated client and where we see that governments do this well is when they succeed in fostering a culture of continuous improvement within the government. So, setting up a system to collect and analyse data and infrastructure project performance, then using that kind of data to identify areas for improvement and best practices. And you can look to the report we just did as kind of a high-level example of this. You're looking at where do we see that costs are low? And where do we see the countries actually have really well-performing infrastructure projects? And what are they doing differently? Doing that within New Zealand I think is pretty important.

Other ideas could be things like establishing clear and measurable performance targets for projects, and then holding developers and contractors accountable for meeting those targets. When we look at infrastructure around the world, we see that this is actually something that clients or governments are not very good at. In Denmark, we have this really large hospital building programme going on and it's nearing completion. And one of the things that happened here was that the Danish government actually did build, I think, 16 large hospitals here. There's been all these news articles about how these projects kept on going over budget, they don't have the quality that was anticipated, we have to go back and change 40 surgery rooms, because the air quality wasn't at standard. And this was really, because the client, the Ministry of Health, failed to actually go in and review and hold the contractors accountable and do QA on the projects. The only thing that was really looked at was really high-level metrics, like cost per square metre and number of square metres built. So, the assumption is made that as long as we're on track with the number of square metres built, then the projects are on track. And they really lost the eye for actual performance. And so in the end, infrastructure was delivered at low cost, and at a reasonable pace, but the benefits are not really manifesting as we were hoping.

A third idea, which kind of relates to that is incentivising good performance. And this is really why I think that it's government that needs to drive behaviour. Because at the moment, we see that, that there's really no incentive for driving good estimating and getting really good performance in projects. We see from government side, that often politicians or project managers, they pressure really low-cost estimates, they pressure getting into the ground quickly on projects because they want to get these big projects started and they want to get their name remembered for these projects, right. At the same time, contractors are really incentivised to produce poor, unrealistic estimates, because they're competing on cost and over time. And so actually implementing a different system that is performance based, and where contractors would be rewarded for good performance – actually have some kind of incentive to produce some good estimates and sticking to the estimates – could really help. So just demanding that drive and incentivising I think is really important.

Peter Nunns: What you're saying there is fascinating, right? Because you're echoing back a lot of the anxieties and worries that we've got about how things are going. This is a quite common story internationally. Your company, Oxford Global Projects, is going in providing advice on infrastructure projects all around the world. Would you be able to tell us a bit more about some of that broader work? And maybe what you're learning from it?

Andreas Leed: Yeah, and I think what we're learning is – it's just so interesting to be a part of – because major infrastructure and these mega projects, it's actually pretty rare that governments take them on. So, when we collect data from them, we look at what do these different countries do? And what's the difference in how they approach infrastructure? And, so, we have a wealth of experience across both geographies and different industries. Central to all we do is that we have this database of project performance that we have collected over the years, and it has data for more than 17,000 projects. We use that data to get insight from historical projects and actually try and apply findings to upcoming projects. Generally, our main goal is setting up major projects for success to help clients overcome what we call the Iron Law of projects. Which is basically the thing that our founder Bent Flyvbjerg, who is the leading researcher in this field, always says "over budget, over time, none of the benefits" over and over again. We see the systematic budget overruns, delays, and benefit shortfalls. And looking at infrastructure over time, we haven't seen any improvement since we started measuring performance by these parameters.

To just – I know I'm rambling a bit – but to include and talk about some of the projects that we worked on, we work on large infrastructure, transport programmes, where we estimate risk using historical data, like, for instance, High Speed Two (HS2), which is Europe's largest infrastructure projects sitting at well over 100 billion pounds at the moment. We've done a ton of benchmarking in the hospital sector, which started with us talking to the Ministry of Health in Iceland, because they were going to build a new hospital there. The only hospital in the country and they wanted us to review their plans. So, these obviously just examples, and we have a lot of different examples of specifics we worked on. But I think for Dirk, and for me, our main focus as data scientists is trying to apply these more data-driven methods to improve projects and particularly estimating. One example that I think is really interesting is the Hong Kong Development Bureau, where we are currently developing an AI-based early warning system for spotting high-risk projects. At the moment, it's kind of a splining project. So we are tracking project cash flow dates, and using machine learning models to predict are these project going to be delayed, and are they going to go over budget?

But there's some quite some quite intuitive and simple behaviours we see in projects. For instance, if projects don't spend cash as quickly as we were thinking they would, then obviously, they could go under budget, but more likely, it's just that they're delayed and they're not progressing as fast as they should be.

Peter Nunns: So, you spend money slower, but you end up spending more of it later on. Because you run into a whole bunch of problems, right?

Andreas Leed: Yeah, exactly. You're going to have a delay and delay is costly as well. And often, actually, delays can be worse for governments, because it's one thing that stakeholders and citizens really easily can track. They know that "oh, we have this new metro system, it's been announced to start working in the summer of 2020. I can't go and catch it, so something happened there. Typical for government." Right? I think that this will be a story that is also typical around the world.

Peter Nunns: There's a flip side to this right. You've talked about the evidence you're finding on bad performance and problems, right. But people are also looking for insight into good performance. Presumably that when the Icelandic folks got in touch with you, they were asking for advice on what could go wrong, but probably also, what could go right. Is that a theme?

Andreas Leed: Yes, exactly. And so that's often also something that we look for, we have quite a quantitative approach to project planning. And so as you've seen with the benchmark report, we're looking to what are the cost drivers, or what are the schedule drivers of projects, but we also look into are specific contract types more efficient, or setting up projects financing in this way does that mean that we have more risk in projects, or it could be one of the things. One of the big things that we're looking at and doing a bit of work into is modularisation. Actually looking at using more modular design, and construction of infrastructure projects, breaking projects down into smaller and more manageable modules. Whether that improves performance, as far as we can tell, this approach is showing to reduce complexity, improve predictability, and it also leads to faster and more efficient project delivery.

Peter Nunns: And that tracks with I think, that theme about smaller scale, standardised projects being more predictable in terms of costs and larger, more complex projects being less so. I mean, I suppose the question I always have with that is, to what extent is that just forced on you by the context in which you're building the project: the geology, the urban environment, the

environmental mitigation and to what extent is it a choice that people are making to do something that's quite bespoke?

Andreas Leed: I think there's always opportunity to modularise, more or less. And we see that everywhere. Even for things like large-scale hospitals, which is basically it's one large piece of infrastructure, which you might not really think is able to be modularised in the same way as solar farms or wind farms that are just kind of modules you can click on. We see technology and efficiency gains. So, I think it can be applied anywhere, you just have to be a bit more creative. And I think the really interesting thing about looking at all this data is that you don't even have to really be that creative. You can just look out into the world and see, what are some clever things that others are doing? And why don't we copy that idea? Right.

Peter Nunns: That's sort of why I guess you do performance benchmarking, right, to know where to look. But I mean, back on the research a little bit, one of the things that I took away from that was the benchmarking is actually very hard. There's a lot of factors that matter that will affect project costs. And it's quite difficult to go and control for all of them. We certainly saw that the work that you guys did, and the analysis – the commentary – we put over the top of it as, as a high-level sort of lay of the land type study, rather than a detailed attempt to drill into all those fine details. Could you talk a little bit about how you approach that issue? And what people seeking to do their own benchmarking might want to consider in the process?

Dirk Pöker: So absolutely, benchmarking I think can certainly be challenging, as there are many factors that can affect cost and schedules. And it can be difficult to find similar projects to create benchmarks and even impossible sometimes. However, even with these challenges, high-level benchmarks are incredibly useful for evaluating infrastructure projects. In our approach, we use statistical analysis to select projects that are as similar as possible, while also ensuring that we have a broad coverage of different types of projects in different countries. What we usually do is we test for known costs, or schedule drivers, such as geographical differences, specific project characteristics, scale. But we do understand that there are many unknown factors that can affect the project cost and the project schedule. This is why the best approach is not to find 100 percent similar projects, because they simply don't exist. But instead, to find projects that are similar, or are of similar type and cost drivers and then consider how your project could compare to this pool of reference projects, right. So, we try to take inspiration from [Daniel] Kahneman, stating that getting more information, even when it's not perfect, still helps you to make better decisions. So, the key is to use high-level benchmarks as a starting point, identify areas for improvement and make more informed decisions. It is important to remember that not all drivers can be measured or predicted and having a distributional understanding can provide valuable insights. Right?

Peter Nunns: Yeah. I guess my simplistic take on it would be if you find something that's at the end of the distribution, you need to go and turn over some rocks, right? And look at what's going on there?

Dirk Pöker: Yeah, but on the other hand, I think one of our main pieces of advice is always to not exclude any distributional information, because projects that go 10 times over budget have happened before. And you're optimistic when you think that this could never happen to you. Right. So it's better to cast the net a bit wider than you normally would and not exclude the outliers actually.

Peter Nunns: Yeah, I guess I was thinking more if I was building a project, and I discovered that it was near the outlier end of the distribution, that would probably trigger me to go and look at what's happening there. Right? And we might look at look at what was happening and find out that actually, there's a whole bunch of good explanations for that. And it's still a project that's being delivered as cheaply as it could possibly given the context. But you would want to ask the

question, right?

Dirk Pöker: Yeah, sure. Sometimes the context is really difficult. For example, I remember one challenging benchmark was we were assessing the feasibility of the first road ever to be built between settlements in Greenland. It's called the Arctic Circle Road. I think Andreas worked a lot on this. I think this was a very good example of how challenging benchmarking can get because it's the first of its type.

Andreas Leed: I think it's a good example of how do you benchmark when there's no data available. Because what we were asked to do is that there's been this road proposal for decades now in Greenland of a road going from Kangerlussuaq and Sisimiut which are two settlements in Greenland. And it's called the Arctic Circle Road, a really nice name. And this road has been discussed for ages and now it's finally getting some more traction, and there's actually a working group looking at and asking "Is it feasible?" – "Is it a good idea to build this road?" – "Do we get any benefits?" – "Is it going to increase tourism?" All that kind of stuff. And they came to us and asked us whether we could look into the risks of building the road. "Are the benefits going to outweigh the costs of building the road?" And so obviously, we have quite a different approach to more detailed estimators. We go out and look for reference projects, similar projects, we do this kind of more almost Bayesian approach where we look at the distribution of costs of similar builds, and then we make an assessment, "Is going to be one of the more expensive or less expensive ones?" Just like you said, Peter. And for Greenland there was no data, there's never been a road built between settlements. But obviously, Greenland has quite unique characteristics, but it's not the only Arctic country. So, we just went out into look into what data could we get from Russia, from northern Scandinavia, from Iceland, from Alaska, Canada, and looked at similar builds. Similar gravel roads, similar constraints in terms of how much drainage is needed and other characteristics that we could actually look at. And we managed to come up with a benchmark of 16 projects, not a big one, but at least 16 projects where we could say, okay, this is the range of costs – you can probably expect something within this range of costs. And what we did is we took the distribution of this range of the distribution of costs and we used it for modelling. So, we modelled that against the benefits estimate and then we also use benefit shortfall data from roads that we simply modelled. What's the likelihood that the benefits are going to outweigh the cost? The Monte Carlo model showed about 65 percent likelihood that benefits were going to outweigh the cost. So that was a way that we were able to quantify the uncertainty in the project, even though there's no other roads in Greenland. And it could be that Greenland has some special characteristics. They might not have as much labour accessible, and there might be some other issues with building in Greenland. But I think, obviously, you can take that into account, you just pick a higher '*p*' level of your estimate of the distribution so that you can actually factor things in more uncertainty.

I think that kind of approach is an example for how you can do benchmarking for something that is really unique. It is also something that could be applied to most other pieces of infrastructure, because what we typically see is that internal project estimates are really good at getting the most accurate case, but they vastly underestimate the worst-case scenario, they also underestimate the best-case scenario. So, we typically see that between the p-50 and p-90 that the most likely worst-case range that is often used in infrastructure budgeting decisions. We see that that's just way too narrow. Typically, the p-90 would be 80 percent, more than doubling the budgets. And sometimes we see that model depth, or you need 30 percent additional budget.

Peter Nunns: We've got a tendency to sort of ask for 10 percent funding contingency, which seems a bit heroic in an environment of 10 percent annual inflation in construction costs. Just briefly, to sort of wrap up, Oxford Global Projects has also done a lot of work looking at cost and schedule overruns for major infrastructure projects, as you've mentioned, and this wasn't the focus of our current research, although you touched on it in your report. But it is a hot topic at the moment. So, what are some key insights from your work in that area?

Dirk Pöker: The main finding, is what we already said before, the so called Iron Law of mega projects, is that major projects systematically have budget overruns and delays, and they don't deliver the benefits that they set out to deliver. This is over and over and over again, it's all over the place, every time. So, research shows that this is mainly due to psychological biases and

to political pressures. And one way to tackle those systematic issues is using more data, like benchmarks and planning, and find new ways to use this data and look at the insights.

Andreas Leed: What we see when looking at these 17,000 projects is that most issues in infrastructure planning are actually planning issues. Projects often they have this tendency to try to explain away and say all these things happen to us – "we didn't realise we couldn't work 24 hours a day in the middle of Copenhagen on our metro system". But really, that kind of risk of being delayed because some of your assumptions are wrong is something that you should be planning for. So, you should be looking at how long is it realistic that it takes to build a metro system in the middle of a major city? Right. And we have some other findings, it's not just related to project planning being bad. There's lots of biases in that, because there's lots of good examples as well.

But some of the other findings are things that we've talked about that what that things like complexity makes it really difficult to plan properly. So, if you can try to remove some of that complexity, you will probably have a more successful project in the end. If your organisation is good, projects should be better, and especially clients evaluating project should be better at handling the uncertainty that is just associated with infrastructure building. We often see that clients just want a single point estimate, and they don't really want to understand uncertainty. And it's often that in most project communication that goes to steering boards we see that it's often certainty that's communicated, it's not the uncertainty or "What are the risks in these estimates?" "What are the assumptions we made?" Because if that was clearly communicated, you wouldn't see examples like the Copenhagen metro trying to work 24 hours a day.

Peter Nunns: And that comes back to what you're saying earlier about the importance of government being a sophisticated client of infrastructure and understanding all the ins and outs of what you're doing and it's likely performance. Fantastic insights on all of that. And thanks so much for taking the time to talk to us Andreas and Dirk. It's been a pleasure working with both of you on this. For those who are interested in learning more, please visit our website to www.**tewaihanga**.govt.nz to find links to both the reports. Thanks again, folks.

Narrator: Thanks for listening to 'Infrastructure for a Better Future'. To find out more about the infrastructure challenges we are facing visit www.**strategy.tewaihanga**.govt.nz.