Exhibit PUB # 58-6

Volume 6 – Board Counsel's Book of Documents

NFAT Review
Subject: Construction Risk Background Literature

INDEX

Tab	Description	Reference
1.	Construction Risk – Background Literature	Flyvbjerg et al., "Delusion and Deception in Large Infrastructure Projects – Two Models for Explaining and Preventing Executive Disaster", *California Management Review*, Vol. 51, No. 2, Winter 2009, pp. 170-193
2.	Construction Risk – Background Literature	Ansar et al., "Should We Build More Large Dams? The Actual Costs of Hydropower Development", *Energy Policy* (2014)
3.	Construction Risk – Background Literature	Hollmann et al. "Variability in Accuracy Ranges: A Case Study in the Canadian Hydropower Industry", 2014 AACE Technical Paper
4.	Principles of Cost Estimating	AACE International Recommended Practive No. 18R-97, "Cost Estimate Classification System – As Applied in Engineering, Procurement and Construction for the Process Industries"

TAB 1

Delusion and Deception in Large Infrastructure Projects

Two Models for Explaining and Preventing Executive Disaster*

Bent Flyvbjerg
Corresponding author
Aalborg University
Dept. of Development and Planning
flyvbjerg@plan.aau.dk
and
Delft University of Technology
Faculty of Technology, Policy, and Management

Massimo Garbuio
University of Western Australia
UWA Business School
massimo@biz.uwa.edu.au

Dan Lovallo
University of Sydney
D.Lovallo@econ.usyd.edu.au

Bent Flyvbjerg, Massimo Garbuio, and Dan Lovallo, "Delusion and Deception in Large Infrastructure Projects: Two Models for Explaining and Preventing Executive Disaster," *California Management Review*, vol. 51, no. 2, Winter 2009, pp. 170-193.
DOI: 10.1225/CMR423
Link to published article:
www.emeraldinsight.com/bibliographic_databases.htm?id=1791300&show=abstract

*The authors would like to thank Sara Beckman and Nuno Gil as well as three anonymous reviewers for their helpful comments.

The Economist recently reported that infrastructure spending is the largest it is ever been as a share of world GDP. With $22 trillion in projected investments over the next ten years in emerging economies alone, the magazine calls it the "biggest investment boom in history."[1] The efficiency of infrastructure planning and execution is therefore particularly important at present. Unfortunately, the private sector, the public sector and private/public sector partnerships have a dismal record of delivering on large infrastructure cost and performance promises. Consider the following typical examples.

In January 2003, Toll Collect – a consortium of DaimlerChrysler, Deutsche Telekom, and Cofiroute of France – was scheduled to start tolling heavy trucks on German motorways for the Federal government. The new tolling system was designed to be a showcase for public-private partnership in infrastructure management. A year later the project was falling apart. The developers had been too optimistic about the software that would run the system. The government was losing toll revenues of €156 million ($244m) a month, caused by delays, and estimated to total €6.5 billion before problems could be fixed. For lack of funds, all new road projects in Germany and related public works were put on hold, threatening 70,000 construction jobs. Politicians and members of the media were calling for prosecution of Toll Collect for deceiving the government. Finally, the German transport minister cancelled the contract with Toll Collect and gave the company two months to come up with a better plan, including how to fill the revenue shortfall. By the time tolling at last started, after further delays in 2005, "Toll Collect" had become a popular byword among Germans used to describe everything wrong with the national economy.

In 1987, Eurotunnel, the private company that operates the tunnel under the English Channel, went public to raise funds for the project. The Channel tunnel was *prime minister Margaret Thatcher's flagship project to show the world how private business could effectively provide public infrastructure. Eurotunnel told investors that building the tunnel--the longest of its kind in Europe--would be relatively straightforward and that 10% "would be a reasonable allowance for the possible impact of unforeseen circumstances on construction costs."*[2] *Once built, the real cost of the project was double the forecasted costs in constant dollars. Initially, the misinformation about costs and risks served the purpose of getting the project started. From the 1987 IPO until cost overruns hit the project one and a half years later, share prices more than tripled. Then they fell by two thirds and, when it became clear that revenue projections were as biased as cost forecasts, by another two-third. In 1995, Eurotunnel stopped interest payments on its loans and began a decade-long, tumultuous process of financial restructuring from which it did not recover until 2007. The intended flagship of privatization became a scare story for business and set back the process of infrastructure privatization by at least a decade.*

In 1959, the construction of Sydney Opera House started before either drawings or funds were fully available. The initial budget of seven million Australian dollars was a political, low-balled budget designed for project approval before the coming elections.[3] *Eventually the Opera House was opened in 1973, 10 years later than the original planned completion date, at a cost of 102 million Australian dollars. It holds the world record for cost overrun at 1,400 percent and this was for a scaled-down version of the original design. This figure does not include 45 million dollars allocated in 2002 in part to bring the building more in agreement with the architect, Jørn Utzon's original plans.*

Over Budget, Over Time, Over and Over Again

There are some phenomena that have no cultural bounds such as maternal love and a healthy fear of large predators. We can add to this list the fact that, across the globe, large infrastructure projects almost invariably arrive late, over-budget and fail to perform up to expectations. Cost overruns and benefit shortfalls of 50 percent are common; cost overruns above 100 percent are not uncommon. For example, in one study of major projects in 20 countries, nine out of ten projects had cost overruns.[4] Similarly, a study of 44 urban rail projects in North America, Europe, and developing nations, including London's Tube and the metros in Washington DC and Mexico City, found that the average construction cost overrun in constant prices was 45 percent; for a quarter of the projects cost overruns were at least 60 percent. In addition, passenger ridership was, on average, 50 percent lower than forecast. Furthermore, for a quarter of the projects, ridership was at least 70 percent lower than estimated.[5] An appropriate slogan seems to be "over budget, over time, over and over again”. As comforting as it is to know that we are not alone in our folly, it would be even better to minimize the gap between expectations and performance for projects that consume such a large share of the private and, especially, public purse.

Executives typically attribute project underperformance to numerous uncertainties such as project complexity, technological uncertainty, demand uncertainty, scope clarity, unexpected geological features and negative plurality (i.e., opposing stakeholder voices).[6] No doubt, all of these factors at one time or another contribute to cost overruns, benefit shortfalls, and time delays. The goal of this article, however, is not to explain, for example, how to implement complex projects more efficiently by over-coming these uncertainties. Rather, we explain why cost, benefits and time forecasts for more complex projects are systematically over-optimistic in the planning phase in comparison to less complex projects. In other words, “why do project planners, on average, fail to anticipate the greater costs of complex projects or those based on new technologies, etc.?”

The underlying reasons for all forecasting errors can usefully be grouped into three categories: 1) delusions or honest mistakes; 2) deceptions or strategic manipulation of information or processes or 3) bad luck.[7] Bad luck or the unfortunate resolution of one of the major project uncertainties is the attribution typically given by management for a poor outcome.[8] While not denying such a salient explanation, this article explores the underlying psychological and governance reasons for mis-estimation rather than proximate engineering causes.

Deliberately or not, risks of scope changes, high complexity, and unexpected geological features are systematically underestimated during project preparation. Both delusion and deception see the high failure rates for ventures as a consequence of flawed decision making. According to the first explanation - delusion - the flaw consists in executives falling victim to what psychologists call the planning fallacy.[9] In its grip, managers make decisions based on delusional optimism rather than on a rational weighting of gains, losses, and probabilities. They overestimate benefits and underestimate costs and time. They involuntarily spin scenarios of success and overlook the potential for mistakes and miscalculations. As a result, managers pursue initiatives that are unlikely to come in on budget or on time, or to ever deliver the expected returns. These biases are often the result of the *inside view* in forecasting: decision makers have a strong tendency to consider problems as unique and thus focus on the particulars of the case at hand when generating solutions.[10] Adopting an *outside view* of the problem has been shown to mitigate delusion. It is applied by ignoring the specific details of the project at hand and uses a broad reference class of similar projects to forecast outcomes for the current project.

According to the second explanation - deception - decision-making is flawed by strategic misrepresentation or the presence of what economists refer to as principal-agent problems. Whereas the first explanation is psychological, the second is due to the different preferences and incentives of the actors in the system.[11] In this situation, politicians, planners or project champions deliberately and strategically overestimate benefits and underestimate costs in order to increase the likelihood that their projects, and not their competition's, gain approval and funding. These actors purposely spin scenarios of success and gloss over the potential for failure. This results in managers promoting ventures that are unlikely to come in on budget or on time, or to deliver the promised benefits. However, this misrepresentation and failure can be moderated by measures that enhance transparency, provide accountability, and align incentives.

While delusion and deception have each been addressed in the management literature before, in what follows they are jointly considered for the first time. In addition, they are specifically applied to infrastructure problems in such a way that both academics from diverse fields and, more importantly, practitioners can understand and implement the suggested corrective procedures. Among its contributions, the article provides a framework for analyzing the relative explanatory power of delusion and deception in such a way that it is possible to disentangle whether non-accurate forecast are more likely to be due to one or the other explanation, or both. Moreover, it suggests a simplified framework for analyzing the complex principal-agent relationships that are involved in the approval and construction of large infrastructure projects. This will facilitate the design of incentive systems and corrective procedures for improving forecasts, which is a major goal of the article.

The article is organized in two main parts. In the first part, we initially discuss the cognitive factors that produce delusion and then political and economic explanations of deception. An analysis of the complementary nature of explanations based on delusion and deception follows. In the second part, we build on these models and show how appropriate financial and non-financial incentives can mitigate deception. To conclude, we illustrate reference class forecasting, an outside view debiasing technique that has proven successful in overcoming both delusion and deception in private and public investment decisions.

Delusion and Deception in Large Capital Projects

Delusion

Our first explanation – delusion – accounts for the cost underestimation and benefit overestimation that occurs when people generate predictions using the inside view. Executives adopt an *inside view* of the problem by focusing tightly on the case at hand, by considering the plan and the obstacles to its completion, by constructing scenarios of future progress, and by extrapolating current trends.[12] In other words, by using typical bottom-up decision making techniques, they think about a problem by bringing to bear all they know about it, with special attention to its unique details. Below we illustrate two cognitive delusions the inside view facilitates: the planning fallacy as well as a heuristic - rule of thumb - called anchoring and adjustment.

When forecasting the outcomes of risky projects, executives often fall victim to the *planning fallacy*. Psychologists have defined it as the tendency to underestimate task-completion times and costs, even knowing that the vast majority of similar tasks have run late or gone over budget.[13] It is a well-established bias in the experimental literature. In one set of experiments, Buehler, Griffin and Ross assessed the accuracy of psychology students' estimates of completion times for their year-long honors thesis project.[14] In the experiments, the students' "realistic" predictions were overly optimistic: 70% took longer than the predicted time, even though the question was asked toward the end of the year.

On average, students took 55 days to complete their thesis, which was 22 days longer than predicted. Similar results have been found with various types of subjects and for a wide variety of tasks such as holiday shopping, filing taxes, and other routine chores.[15]

These findings are not limited to the laboratory. Cost and time overruns in large infrastructure projects have been studied by considering numerous contractual arrangements. In the case of conventional procurement, in which the public entity separately engages with several private companies, each of them providing a specific part of the service, costs and times overruns have been systematically observed in a wide range of projects.[16] In business, executives and entrepreneurs seem to be highly susceptible to this bias. Studies that compared the actual outcomes of capital investment projects, mergers and acquisitions, and market entries with managers' original expectations for those ventures show a strong tendency towards overoptimism.[17] An analysis of start-up ventures in a wide range of industries found that more than 80% failed to achieve their market-share target.[18]

Anchoring and adjustment is another consequence of the inside view thinking that leads to optimistic forecasts.[19] Anchoring on plans is one of the most robust biases of judgment. The first number that is considered as a possible answer to a question serves as an "anchor". Even when people know that the anchor is too high or too low, their adjustments away from it are almost always insufficient. A classic experiment revealed the power of anchoring and insufficient adjustment. People were asked to estimate various percentages, such as the percentage of African countries in the United Nations.[20] For each quantity, a number was determined by spinning a wheel of fortune in the presence of the subject. The subjects were first asked to indicate whether the number was higher or lower than the percentage of African countries and then to estimate the percentage by moving upward or downward from the arbitrary number. The arbitrary number had a substantial effect on the estimates. For instance, the median estimate of the percentage of African countries in the United Nation was strongly related to the starting points: individuals who received 10, estimated 25% whereas those that received 65, estimated 45%. These subjects started from a random anchor and then insufficiently adjusted away from it.

Similar results have been found with experienced real estate brokers who were asked to assess the value of a property.[21] These agents unanimously agreed that they did not factor a house's listing price into their evaluation of its "true" value. Each of the agents was given a 10 page booklet on the house that was being sold, which included information specific to the house as well as information about the prices and characteristics of other houses in the area that had recently been sold. The only difference in the information that the various brokers received was the listing price of the house, which was randomly manipulated within a range of plus or minus 11% of the actual listing price. The agents then went out and visited the house that was being sold, as well as several other houses in the neighborhood. The listing price significantly affected these experienced agents' evaluations. Furthermore, when told about the results, the agents maintained that the listing price anchor had no effect!

In the context of planning for a large infrastructure project there is always a plan, which is very likely to serve as an anchor. Furthermore, the plan that is developed is almost always seen as a "realistic" best or most likely case. Executives know that events may develop beyond the best or most likely case so they generally attempt to capture unforeseen costs by building in a contingency fund that is proportional to the size of the project (e.g. for cost overruns in capital investment projects). However, when compared with actual cost overruns, such adjustments are clearly and significantly inadequate.[22] Furthermore, the initial estimate serves as an anchor for later stage estimates, which never sufficiently adjust to the reality of the project's performance.

The power of these heuristics and biases is well illustrated in a field study where the Rand Corporation examined 44 chemical process plants (Pioneer Process Plants), owned by firms such as 3M, du Pont and Texaco among others. Actual construction costs were over twice as large as the initial estimates.[23] Furthermore, even a year after start-up about half of the plants (21) produced at less than 75% of their design capacity, with a quarter of the plants producing at less than 50% of their design capacity. Many of the plants in this latter category had their performance expectations permanently lowered. As illustrated in Figure 1, the typical initial estimate is less than half the final cost. Furthermore, at every subsequent stage of the process, managers underestimate the cost of completing the construction of Pioneer Process Plants.

Add Figure 1 here

Deception

Our second explanatory model – deception – accounts for flawed planning in decision making in terms of politics and agency issues. With this model we introduce political and organizational pressures in executive decision-making. We describe the principal-agent problem first and the sources of strategic deception second.

The Principal-Agent Problem in Large Capital Investments

In this section we talk about principal-agent (P-A) problems, which have mainly been examined in the context of private firms but which can be even more pernicious in public situations.[24] These are defined by relationships where a principal engages an agent to act on his or her behalf. Typical examples include a Board hiring a CEO to manage the company on behalf of the shareholders or a manager hiring an employee to carry out tasks. In fact, there is a P-A relationship for every two levels in an organization.

Large capital investment projects are situations where a multi-tier principal-agent problem exists.[25] An example will help illustrate the framework. Consider a local government that intends to build a new tunnel across a large capital city, for the benefits of the local residents and, more broadly, of the state population. The focal project will compete with other projects for funds from the state government. Once the approval is obtained, the local government puts construction out for tender. The winning bidder will carry out the construction of the infrastructure.

Figure 2 graphically represents the complexity of the P-A relationships in the case of a large capital investment proposed by a local government to the state government. In this specific example, there are three tiers of P-A relationships.

Add Figure 2 here

The first tier encompasses the relationship between taxpayers and the state government. Taxpayers are the principal, whereas the state government is the agent of the taxpayers that is supposed to act in their interest. As the final beneficiaries of the infrastructure, taxpayers expect projects to deliver the largest possible benefits to the community, by incurring minimal costs, attenuating risks and reaching completion within an agreed timeline. Individuals in the state government, who are elected by taxpayers, typically have their own interests, for example being re-elected and/or being remembered for the building of monumental infrastructures.

The second tier of P-A relationships has the local government acting as the agent of both taxpayers and state government. With respect to the taxpayers, the local government has the duty to propose infrastructures that provide the largest benefits to the community, and that are delivered on budget and on time. With respect to the state government, it has a duty of suggesting the best allocation of the taxpayer funds. Moreover, holding the most complete data about costs and benefits of the infrastructure that it proposes, it has the duty of providing the state government with the most accurate forecasts needed to make an informed decision. However, given the competition for scarce resources, the local government has an interest in understating its risks and costs, while overstating their benefits.

The mechanism of benefit overestimation is very simple as explained by an interviewee in research done by one of us (Flyvbjerg) and the Danish consultancy firm Cowi: “The system encourages people to focus on the benefits--because until now there has not been much focus on the quality of risk analysis and the robustness [of projects]. It is therefore important for project promoters to demonstrate all the benefits […].”[26] In addition, knowing that the next election usually happens before the time that the proposed project is built and sometimes even approved, the local government has little interest in providing accurate forecasts. As shown by Flyvbjerg in his research on the Sydney Opera House, Joe Cahill, the Labor-premier of New South Wales, publicized a political budget for approval and fast-tracked construction to start before the elections, in case the Labor party lost the elections and attempts would be made to stop the project. According to Bob Carr, who later followed Cahill as premier of New South Wales, Cahill instructed his people "to go down to Bennelong Point [the site of the Sydney Opera House] and make such progress that no-one who succeeds me can stop this [the Opera House] going through to completion."[27] When the premier's lowballing and fast-tracking of the Opera House inevitably led to cost overruns, the architect, Jørn Utzon, was blamed. Utzon prefers to remain out of the public eye, but his son, architect Kim Utzon, explains in lieu of his father: “It was a political decision to publicize a low budget for the building, which was expected to gain approval in the political system, but which was very quickly exceeded. So even if the cost overrun turned out to be 1,400 percent in relation to the publicized budget, this budget was an eighth of the real budget for the building. So the real cost overrun is only 100 percent. The rest was politics.”[28]

The third tier of P-A relationships involves the local government as the principal of agents hired to provide specific services, such as analysts and planners as well as contractors. Analysts and planners are engaged to gather the information necessary for making the final go/no-go decision. They have an incentive to provide information that is compatible with pleasing the local government, having the project approved, and being re-engaged on the next project. A manager on a large infrastructure project explained to Flyvbjerg and Cowi in their research on transport infrastructure management in the UK: "Most decent consultants will write off obviously bad projects but there is a grey zone and I think many consultants in reality have an incentive to try to prolong the life of projects, which means to get them through the business case. It is in line with their need to make a profit."[29] Another interviewee in the same study recognized that planners have better information than politicians but have no incentive to reveal such information, but rather the opposite: “You will often as a planner know the real costs. You know that the budget is too low but it is difficult to pass such a message to the counsellors [politicians] and the private actors. They know that high costs reduce the chances of national funding.” Similarly, builders have the primary interests to win the tender, by offering the lowest possible price, since they know that re-contracting is often possible and, unless the contract is a fixed price and lump sum contract, delays will be tolerated. Even if interests are divergent in this case, cost-overruns and delays are tolerated unless the local government is held responsible. Clearly, the multi-tier relationship described in this specific example of the construction of a cross-city tunnel can be easily extended to the approval of any kind of large public infrastructure project.

Sources of strategic deception

There are certain conditions, however, that make strategic deception more likely within *each* P-A relationship. Self-interest, asymmetric information, differences in risk preferences and time horizons as well as the clarity of accountability are among the most cited causes.

A necessary condition for P-A conflicts is a difference in the actors' *self-interest*. Executive ventures, public and private, are often multimillion- and sometimes even multibillion-dollar projects. When they go forward, many stakeholders (e.g. contractors, engineers, architects, bankers, landowners, construction workers, lawyers, accountants and developers) have widely divergent incentives. In addition, politicians and executives may use ventures to jockey for position and to build monuments, which allows administrators to get larger budgets, and cities to acquire investments in infrastructure that would otherwise go elsewhere.[30] If these stakeholders are involved in, or indirectly influence, the forecasting of costs and benefits at the approval stage (the business case), this is liable to bias the entire subsequent process.

Political and economic self-interest also exists at the level of cities and states. Pickrell pointed out that transit capital investment projects in the US compete for discretionary grants from a limited federal budget each year, and that this creates an incentive for cities to make their projects look better, or else some other city may get the money.[31] Flyvbjerg and Cowi found similar results for the UK.[32] Altshuler and Luberoff, Delaney and Eckstein, and Morris and Hough found corresponding results for other project types, including major roads, tunnels and bridges, airports, stadiums, power stations, oil and gas extraction, and IT systems.[33]

A second source of strategic deception is the presence of *asymmetric information*, which means the agent who champions a project (e.g. the local government in the example above) has information that the principal does not (e.g. the state government). Being unaware of all the relevant information, the principal and ultimate decision maker may be easy to deceive. In a recent study, Flyvbjerg and Cowi interviewed public officials, planners, and consultants who had been involved in the development of large UK transportation infrastructure projects.[34] This study shows that strong interests and strong incentives exist at the project approval stage to present projects' costs and benefits as favorably as possible. Local authorities, local politicians, local officials, and some consultants (as agent's) all stand to benefit from a project that looks favorable on paper and have little incentive to actively avoid distorted estimates of benefits, costs, and risks. National bodies, like the Department for Transport and the Treasury, act as a principal and usually fund and oversee projects. They usually have an interest in more balanced appraisals, but so far they have had little success in achieving such balance. This situation may be changing with recent schemes to curb the optimism bias, which were initiated by HM Treasury in order to gain better predictability and control in public budgeting.[35]

A third relevant source of P-A issues is the presence of *different risk preferences*. For instance, if the principal is risk averse, the agent who submits a proposal for approval will have to downplay the possible risks of the venture in order to convince the principal. Because of this, managers lower down in the organization may have to misrepresent, hide and manipulate information in order to get requests for funds approved.

In addition to any inherent difference in risk preferences between decision makers and organizations or the public, another factor that comes into play is the vastly *different time horizons* the actors use to evaluate the decision. Typically, this is very long for taxpayers, but less than a decade for the individuals that are acting on their behalf. These agents may also be concerned with being remembered for initiating monumental infrastructure or, more prosaically, being re-elected. In fact, whereas a standard election term is 4 years, the average length of the time from the start of planning to start of operations for a large infrastructure project is commonly 10-15 years.[36]

Finally, another condition that leads to strategic deception is *diffuse or asymmetric accountability*. When multiple people are responsible for the ultimate success or failure of a project, it can be difficult for any one agent to be held accountable for a bad outcome. If a new initiative fails, it is often hard to place the blame squarely on one or a few actors. This lack of accountability *ex-post* can cause the agent to promote *ex-ante* projects that protect them from being held accountable if it fails, which may not be the project that maximizes the principal's total payoff.[37] In addition, the lack of clear accountability can exacerbate the problem of asymmetric information and differences in risk preferences, causing the agent to take more risk than the principal would like.

Diagnosing the Relative Impact of Delusion and Deception

Delusion and deception are two complementary rather than alternative explanations of failure of large infrastructure projects due to cost underestimation and benefit overestimation. Although, in practice, it is often difficult clearly disentangle the two explanations. There are situations, however, where the explanatory power of one of the two models is relatively higher.

The relative strength of each explanation depends on different factors. The key to minimizing delusion is to have a good learning environment. Learning occurs "when closely similar problems are frequently encountered, especially if the outcomes of decisions are quickly known and provide unequivocal feedback".[38] Whereas, the problem of strategic deception occurs when incentives are mis-aligned. The underlying causes of these mis-alignments are differences in preferences, time horizons, incentives, and information between principals and agents.

Figure 3 describes situations where we can expect explanations due to delusion, deception, both, or neither to operate. The figure is divided in four regions. When the learning environment is good and incentives are well-aligned, there is minimal scope for delusion or deception and forecasts tend to be unbiased. Weather forecasts are an example. Meteorologists have no reason to lie and the feedback they receive is so frequent and unambiguous detailed computer models guide their predictions.

Small entrepreneurs' who own the vast majority of their companies have incentives that are well aligned so that we expect most of their errors to be due to delusions. These forecasting errors tend to be quite large. For example, 33% of entrepreneurs perceive their chances of success to be certain, which is obviously is deluded given that over 80% of such ventures fail.[39]

Many computer gaming companies release numerous titles frequently, so their learning environments are good yet they continuously state release dates they do not stick to. This type of deception has been labeled "Cheap Talk" and is designed to pre-empt sales of competitors' products.[40]

The largest errors occur when delusions and deception operate simultaneously. Even within this section useful distinctions can be made. For example, in the private sphere, we consider the construction of process plants. These need to be distinguished between normal process plants and pioneer process plants. They both have similar incentives since they are owned by the same companies, but the former is relatively more frequent and therefore learning is improved and forecasts are more accurate.

In the public sphere, we consider the examples of rails and roads. Cost underestimation and overrun for rail are on average approximately twice that for roads.[41] Similarly, average ridership overestimation for planned rail projects is around 100%, whereas such bias is not found for road projects.[42] The differences between rail and road are statistically highly significant and may largely be explained in terms of differences in incentive structures and the possibility to learn from previous and similar projects. In fact, rail projects typically compete for discretionary grants from a limited national or federal budget. This creates an incentive for promoters (agents) to make their projects look better on paper with artificially high benefit/cost ratios, or else the central government (the principal) may decide to fund some other project.[43] That is, there are incentives to provide biased estimates.

In the case of roads, funds are typically allocated as block grants with a certain amount of dollars allocated for road building as where individual projects do not compete for funds directly against each other or against other types of projects outside the highway agency. As a consequence, the misalignment of incentives between promoters and approvers are higher for rail than for roads. This conclusion is supported by a study of stated causes for inaccuracies in traffic forecasts for 234 rail and road projects. For rail projects, deception in terms of "deliberately slanted forecasts" was explicitly stated as a main cause of inaccurate (biased) forecasts in 25% of projects, whereas this was the case for zero road projects.[44] This does not mean that estimates of costs and benefits of planned roads are never deceptive. This source of bias, however, appears to be less prevalent and less systematic for road than for rail projects. Finally, road projects are more common than rail projects, so the opportunity for learning is greater as well.[45]

For the largest public infrastructure projects, such as the Sydney Opera House, the Channel tunnel, concert halls or stadiums, the cost blowouts appear to be more heavily weighted towards deception. This suggests new governance procedures are needed to minimize waste in the upcoming infrastructure boom.

------- ------ --------

Add Figure 3 here

------ ------- --------

Overcoming Delusion and Deception in Large Infrastructure Projects

Delusion and deception are not insurmountable. While every large infrastructure project has its own idiosyncrasies, these projects are all prone to delusion and deception, regardless of the private or public funding institution. This section of the article addresses some prescriptive governance advice for overcoming delusion and deception. First, we focus on possible techniques useful for overcoming deception. We address financial and non-financial incentives[46] for the agents, namely the proposing and the approving institutions, planners and contractors. Then, we discuss the adoption of an outside view, which deliberately avoids the details of the case at hand and simply focuses on understanding the historical statistics and patterns of similar projects. The specific outside view tool we use is reference class forecasting, where decision makers search for an unbiased, representative population of similar and past cases that will become useful to make unbiased predictions of the future. When delusion and deception are difficult to disentangle, reference class forecasting has been successful in overcoming both dysfunctional behaviors in diverse settings, including major transport scheme forecasting[47] and movie forecasting.[48]

Overcoming Deception through Accountability and Transparency

Financial and non-financial incentives should be given to the agents of taxpayers and the institution proposing the project. We first discuss incentives for proposing and approving institutions and then we turn our attention to the incentives for planners and bidders.

Incentives for the Agents of the Taxpayers: Institutions Proposing and Approving Large Capital Infrastructure Projects

As shown by Flyvbjerg,[49] artificially low costs, exaggerated benefits and underestimated risks are common strategies employed by the proposing institution to have a large infrastructure project approved. This is facilitated by the asymmetric information existing between those who propose a large infrastructure project and those who approve and fund it. In addition, lack of clear accountability and the misalignment of time horizons may lead the proposing individual(s) to take more risk than the funding institution or the taxpayers would like. To overcome deception (as well as empire building motives) there are two key best practices that have been employed: (1) the proposing and the approving institutions should share financial responsibility and (2) private financers should participate in financing the project with their own capital at risk.[50] We examine these issues in turn.

Institutions proposing and approving large infrastructure projects should *share financial responsibility* for covering cost overruns and benefit shortfalls resulting from misrepresentation and bias in forecasting, which helps align incentives. In a recent consultation document,[51] the U.K. Department for Transport proposes a requirement for all large infrastructure projects that asks for funds from the Department to have a minimum local contribution of 10% (25% for light rail) of the gross cost in order to gain program entry, upon the belief that "if an authority has a financial stake in a scheme this provides a clear incentive to ensure that the right structures and resources are in place to bring it to fruition to time and budget".[52] Recognizing that planners are subject to optimism bias, the Department for Transport requires that all requests for funds include an "Optimism Bias Uplift", which is an empirically based adjustment to a project's costs for different percentiles of cost overruns, on the basis of the project type. The uplifts are computed on the basis of actual cost overruns in a reference class of completed projects comparable to the project seeking funding. For example, if the funding institution was prepared to accept a 50% risk of cost overrun (the '50% percentile') on a road scheme then an uplift of 15% should be applied. But if the funding institution was only prepared to accept a 20% risk of cost overrun (the ‘80% percentile’) then a higher uplift of 32% was recommended. In addition, those uplifts are different for project types - the same 50% percentile uplift for rail would be 40% (which is almost three times the amount for roads) – and for different stages of development – uplifts are the highest for “programme entry” and the lowest when the project is receiving “full approval”. Notice that these top down estimate uplifts encompass the different complexities of the projects without going into specific details, which is the point of the outside view.

The “Optimism Bias Uplift” is useful to control cost underestimation before the approval of projects. If no measure is taken to control cost escalation after project approval, promoters would simply “postpone” the appearance of costs during the project construction. Therefore, to discourage cost increases during the implementation of the project, the Department for Transport advises that requests for funds should include an additional risk allowance in the amount of 50% of the Optimism Bias uplift. If during the implementation phase a project requires further expenditures that are within the risk allowance, these do not require additional approval from the funding institution. However, the local authority is expected to contribute at least 50% of the cost increase. The local authority should be expected to fund any expenditure in excess of that risk allowance.[53] In this way a clear incentive has been imposed on the local authority to avoid cost overruns. Such an incentive did not exist before.

Beside the provision of the Optimism Bias uplifts, the UK Department for Transport requires promoters to construct a comprehensive Risk Register to mitigate the risk involved in the implementation of large schemes. This register lists the risks that are likely to affect the delivery and operation of the proposed infrastructure. Construction risks (e.g., timescale and cost perspectives) and operational risks (e.g., maintenance risk and revenue risk) and a share of risks associated with climate change should be included in the register. In addition, it advises that the Risk Register “needs to identify who owns the identified risk. For example, some risks may be transferable through insurance or financial instruments”.[54]

In addition, to obtain more realistic forecasts and reduced risk, full public financing or full financing with a sovereign guarantee should be avoided. Whenever possible, the decision to go ahead with a project should be made contingent on the willingness of private financiers to participate *without a sovereign guarantee for at least one third of the total capital needs*. The lower limit of a one-third share of private risk capital is based on practical experience.[55] Contracts should be written in such a manner that risk allocation is balanced, i.e., the risk to private financiers must be real, with no comfy escape clauses that return risk to the taxpayer when things get difficult. This has been beneficial in situations where the project passed the market test as well as when it did not (i.e., whether the project has been subsidized or not). Private lenders, shareholders, and stock market analysts should produce their own forecasts or should critically monitor existing ones. If they were wrong about the forecasts, they and their organizations should be held responsible for the mistakes.

Incentives for the Agents of the Proposing Institution: Planners and Bidders

Incentives for planners and bidders encompass two measures that address strategic misrepresentation in the proposal and the bidding phases respectively. We discuss them in turn.

To decrease the likelihood of strategic misrepresentation of costs, timeframe and benefits, incentives aiming to achieve higher transparency should be used in order to align the incentives of the planners (and the proposing institution) to provide more accurate forecasts. To provide incentives to planners to disclose their information regarding the specifics of the projects, *rewards* and *higher criticisms* of the forecasts could be two fruitful alternatives. For instance, financial and non-financial rewards should be promoted for planners who provide realistic estimates. In addition, forecasts should be subject to detailed assessment and criticism, such as expert and independent peer reviews, for projects with major public funding carried out by national or state accounting and auditing offices, like the General Accounting Office in the US or the National Audit Office in the UK, and for projects with private funding by independent private auditors. Other forms of scrutiny, such as public hearings and presentations of the forecasts to the scientific community, should also be encouraged. In the most egregious instances, criminal penalties for seriously misleading forecasts may be warranted.[56]

However, these measures are not sufficient because they do not address opportunistic behaviors that are possible in the bidding phase. In the case of the Sydney Opera House, the Sydney Opera House Act was approved in 1960 with the provision that every 10% increase in the budget would require the Act to be amended by Parliament. This did not hold promoters, bidders and contractors from supporting underestimated costs and time for completion. In fact, during the tender, bidders can act opportunistically by assessing the probability that compensation is possible after the construction stage has been initiated. If compensation is possible, bidders will bid the lowest possible value in order to win the tender. The winning bidder will be typically the bidder who most underestimates the true costs of the project. We call this, the “winner’s blessing”.[57] After the project bas been initiated, the initial low price will be compensated through overpricing the expected scope increases, which the experienced bidders know are almost certain. When compensation is not possible, there is less chance that the bidding price is artificially low.

In this situation, the incentive is to *place financial risk with bidders*. In doing so, bidders have the incentive to disclose any specific information they have regarding costs and completion times, which is often a source of asymmetric information between bidders and the institution managing the construction of the infrastructure. If appropriate measures to overcome deception are taken in the bidding stage, the construction stage should go rather smoothly. However, *placing financial risk with contractors* for delays and scope increases should be a safeguard to be used in conjunction with other measures, especially with regard to coordination between contractors and between contractors and the client.

So far, we have illustrated several contractual arrangements that can be used to decrease the likelihood of strategic misrepresentation of costs, timeframe and benefits by transferring risks to the entities that would otherwise benefit from such misrepresentations. Rather than being part of separated arrangements, these measures can be *bundled* into one contract and a private sector entity can be charged with providing a flow of infrastructure services over time that goes beyond the provision of the building. For example, with a Design-Build-Finance-Operate-Maintain (DBFOM) contract,[58] the private sector entity is responsible for the design, building, financing, operation and maintenance of an infrastructure under a very long period of time, usually 20-30 years, after which the facility is transferred to the public entity. This type of contract addresses in one place several of the mis-alignment issues discussed in this article. These are relatively new mechanisms and deserve further attention.

Table 1 summarizes the key causes of deception and the proposed prescriptive advice.

------- ------ --------

Add Table 1 here

------ ------- --------

Overcoming Delusion and Deception through the "Outside View"

There are several instances where delusion and deception cannot be disentangled. To overcome both of them, the key recommendation is to adopt the *outside view* and, in particular, a forecasting method called "reference class forecasting".[59]

The inside view is the conventional and intuitive approach in planning new projects. The traditional way to think about a project is to focus on the project itself and its details, paying special attention to its unique or unusual characteristics, and trying to predict the events (e.g. strikes or weather) that could influence its success. It essentially ignores the details of the case at hand, and involves no attempt at detailed forecasting of the future history of the project. Instead, it focuses on the statistics of a class of cases chosen to be similar in relevant respects to the present one. For example, similarity could be determined by project type, governance structure, complexity, etc. The case at hand is also compared to other members of the class, in attempt to assess its position in the distribution of outcomes for the class.[60] Using the outside view executives and forecasters are not required to make scenarios, imagine events, or gauge their own and others' levels of ability and control, so they do not risk mis-estimating these factors.

When both the inside and the outside view of forecasting are applied with equal skill, the outside view is much more likely to produce a realistic estimate.[61] In very few instances, since it is based on historical precedent, the outside view may fail to predict extreme outcomes such as those that lie outside all historical precedents. But for most projects, the outside view will produce more accurate results.

The outside view can be implemented through a technique that has been defined as reference class forecasting. It requires the decision maker to obtain a reference class of past, comparable cases when making predictions about costs and benefits of a new project. By introducing distributional information of successful as well as unsuccessful past projects, the decision maker is forced to consider the entire distribution of possible outcomes. This prevents the decision maker from focusing on easily recalled similar projects, which are typically successful ones.[62] The implementation of reference class forecasting is organized into five steps.[63]

(1) **Select a reference class**. Identifying the right reference class involves both art and science. The decision maker usually has to weigh similarities and differences on many variables and determine which are the most meaningful in judging how the project at hand will play out. Sometimes that is easy. A planner who has to forecast the construction costs of a rail project planned in the manner that such a project is usually planned around the world would easily find a reference class. In other cases, especially when the project requires the incorporation of a new technology, it is more difficult. The key is to choose a class that is broad enough to be statistically meaningful but narrow enough to be truly comparable to the project at hand.

(2) **Assess the distribution of outcomes**. Once the reference class is chosen, the decision maker has to document the outcomes – in terms of whichever variable is considered pertinent, e.g. cost overrun, total costs or unit costs – of the prior projects and arrange them along a distribution of outcomes, showing the extremes, the median, and any clusters. Sometimes it will not be possible to precisely document the outcomes of every member of the class. However, a rough distribution can still be obtained by calculating the average outcome as well as a measure of variability. Obtaining good projects with valid data is the hardest and most time-consuming part of reference class forecasting. Results will only be as good as the projects and data that are used as input to the exercise. As part of data validation it must be decided what type of issues data take into account, e.g., uncertainty, design changes, and the like.

(3) **Make an intuitive prediction of your project's position in the distribution**. Based on his or her own understanding of the project at hand and how it compares with the projects in the reference class, the decision maker needs to predict where it would fall along the distribution. Because the intuitive estimate will likely be biased, the final two steps are intended to adjust the estimate in order to arrive at a more accurate forecast.

(4) **Assess the reliability of your prediction**. This step is intended to gauge the reliability of the forecast made in Step 3. The goal is to estimate the correlation between the forecast and the actual outcome, expressed as a coefficient between 0 and 1, where 0 indicates no correlation and 1 indicates complete correlation. In the best case, information will be available on how well the decision maker's past predictions matched the actual outcomes. Then, the correlation based on historical precedent can be estimated. In the absence of such information, assessments of predictability become more subjective. An estimate of predictability may be based on how the situation at hand compares with other forecasting situations. Through a diligent statistical analysis, the decision maker could construct a rough scale of predictability based on computed correlations between predictions and outcomes for other endeavors like road or bridge construction. He or she can then estimate where his or her ability to predict rail project construction costs lies on this scale. When the calculations are complex, it may help to bring in a skilled statistician.

(5) **Correct the intuitive estimates.** Due to the bias, the intuitive estimate made in Step 3 will likely be optimistic – deviating too far from the average outcome of the reference class. In this final step, the estimate is adjusted toward the average based on the analysis of predictability in Step 4. The less reliable the prediction, the more estimate needs to be regressed toward the mean. Suppose that the intuitive prediction of the construction costs is $4 billion and that, on average, rail

projects in the reference class cost \$7 billion. Suppose further that the correlation coefficient has been estimated to be 0.6. The regressed estimate of construction costs would be:

$$\$7B + [0.6 (\$4B - \$7B)] = \$5.2B$$

Thus, the adjustment for optimism will be substantial, particularly in highly uncertain situations where predictions are unreliable.

Besides overcoming optimism bias and mitigating the problems that derive from strategic misrepresentation, reference class forecasting provides two other major benefits. First, since the reference class comprises previous projects, it helps to look across types of projects, geographies and types of financing methods to provide much more information to the decision maker. This represents a test to see which projects have worked out in the past and the kind of tweaks to the model that might increase the chance that the project at hand would be a success. Second, it helps to provide a reality check on whether the project is likely to perform up to expectations.

The first instance of reference class forecasting in practice was carried out in 2004 estimating the costs of the proposed Line 2 of the Edinburgh Tram. The business case estimations prepared by the promoters included a base cost of £255 million (US\$400 million) and an allowance for covering contingency and optimism bias that amounted to 25% above the base cost. However, a reference class of 46 comparable rail projects indicated that the promoters' estimates were optimistic and that optimism bias uplifts should be applied. The reference class was established by Bent Flyvbjerg and Cowi in close collaboration with experts at the UK Department for Transport.[64] Rail projects from Flyvbjerg's megaproject database were screened for inclusion, focusing on projects that had been planned and built under comparable regulatory and contractual regimes. Statistical tests were applied to decide whether projects were indeed comparable, before inclusion in the reference class. After this, an independent review applied optimism bias uplifts to the promoters' capital cost estimates, based on cost overruns in the reference class as required by the UK Department for Transport, and thereby taking an outside view on the cost forecasts. The review concluded that total capital costs were more likely to be £357 million, with a 50% risk of going over budget. If the client wanted to reduce the risk of going over budget to 20%, higher uplifts had to be used and the capital cost budget would have been £400 million.[65]

The resulting reaction to the reference class forecast was that the proposed tram was re-evaluated. Scottish Finance Secretary John Swinney publicly stated, "I want to be absolutely sure about the calculation of the costs involved in these projects, and the assessment of risk involved, before they progress further" (*The Scotsman*, June 6, 2007).

The benefits of using reference class forecasting are present both when delusion is substantial (and deception is relatively less important) and when deception is substantial (and delusion is relatively less relevant). The key is to choose a similar class of situations to the focal problem. In the first instance, cost and benefit estimates are affected by optimism. Reference class forecasting bypasses bottom up estimates for the project at hand and uses the actual outcomes (not biased estimates) of similar past projects. In the second instance, these estimates are not affected by optimism but have been purposely misrepresented to get the project approved. By using realized outcomes of similar past projects rather than manipulated estimates of the current project, reference class forecasting provides more reliable, top down estimates of the true costs and benefits of the project. Reference class forecasting helps both to avoid common cognitive biases and strategic manipulation in order to produce more accurate forecasts. A new and related forecasting technique called Similarity Based Forecasting may provide even more accurate forecasts but has yet to be proved on infrastructure projects.[66]

The Way to Go

Large infrastructure investments are a vital component of any public or private institution. Unfortunately, cost overruns, delays and exaggerated benefits are the norm rather than the exception for roads, bridges, stadiums, concert halls, new plants, etc.

Although large infrastructure projects occur frequently across the globe, any individual project is often a once in a career decision for a public or private executive. Thus it is difficult for executives to learn from their own prior mistakes. It is rare for executives to deliberately learn from similar projects other have attempted. Typically executives to adopt an *inside view* of any particular problem – where they focus on the specifics of the case at hand. Without the opportunity to learn from rapid and unambiguous feedback regarding their estimates of costs and benefits, executives can hardly learn from these unique decisions and avoid making similar mistakes in future projects. In such situations, inside view thinking leads to numerous cognitive biases that result in optimistic delusions.

These, often individual, optimistic delusions are confounded, sometimes even dwarfed, by the magnitude of strategic deceptions among the different actors in the system. On several occasions, however, decision makers have attempted to justify their deceptive behavior by arguing that the decision was in the *public interest*. On one hand, it can be argued that public-sector executives may decide to deliberately underestimate costs in order to provide public officials with an incentive to cut costs and thereby to save the public's money. According to this type of explanation, higher cost estimates would be an incentive for wasteful contractors to spend more of the taxpayer's money. Empirical studies have identified executives and planners who say they deliberately underestimate costs in this manner to save public money.[67] Merewitz endorsed and summarized this viewpoint as "keeping costs low is more important than estimating costs correctly".[68]

On the other hand, a second explanation in terms of public interest covers the not uncommon situation where project promoters believe their venture will benefit society and posterity. They feel that they should do anything possible to make the project happen, including cooking forecasts of costs and benefits. Both types of public-interest explanations see the end (project approval) as justifying the means (estimates of costs and benefits that show the project should be approved).[69]

However, these arguments overlook an important fact. Underestimating the costs and overestimating the benefits of a given project results in artificially high benefit-cost ratio, which in turn leads to two problems. First, the project may be started despite the fact that it is not economically viable. Second, a project may be started instead of another project that would have yielded higher returns had the actual costs and benefits of both projects been known. Thus, for reasons of economic efficiency alone, the argument that cost underestimation saves money must be rejected.[70] As a case in point, an ex post benefit-cost analysis of the Channel tunnel between France and the UK showed that the actual net present value of the project to the British economy was minus US$17.8 billion and the actual internal rate of return minus 14.45 percent. The study concluded that "The British Economy would have been better off had the Tunnel never been constructed".[71]

Because delusion is often accompanied by strategic deception, this study's prescriptive advice has been broken into two parts. First, we focused on best practices to diminish strategic deceptions (e.g. P-A issues) in the specific context of infrastructure projects. Next, we examined how executives can adopt an "outside view" of problems by using reference class forecasting. This statistical procedure uses both a forecaster's intuition and historical data to mitigate the two types of errors and arrive at a more accurate estimate. The American Planning Association has recommended this procedure for large infrastructure projects. Its widespread use would surely produce more accurate estimates of large infrastructure projects and projects like Toll Collect and the Channel tunnel would be profitably and happily foregone by the vast majority of the public. Ultimately, accurate reference class forecasting, proper incentives and budgets are the way to go.

Notes

---

[1] "Building BRICs of growth," *The Economist,* June 7, 2008, p. 80.

[2] "Under Water Over Budget," *The Economist*, October 7, 1989, p. 37–38.

[3] B. Flyvbjerg, "Design by Deception: The Politics of Megaproject Approval," *Harvard Design Magazine,* 22 (2005): 50-59.

[4] B. Flyvbjerg, M.S. Holm, and S. Buhl, "Underestimating Costs in Public Works Projects: Error or Lie?," *Journal of the American Planning Association*, 68/3 (Summer 2002): 279-295.

[5] B. Flyvbjerg, "Cost Overruns and Demand Shortfalls in Urban Rail and Other Infrastructure," *Transportation Planning and Technology*, 30/1 (February 2007): 9-30.

[6] JD. Ford, "The Effects of Causal Attribution on Decision Makers' Responses to Performance Downturns," *Academy of Management Review,* 10/4 (1985): 770-786; S.E. Clapham and C.R. Schwenk, "Self-Serving Attributions, Managerial Cognition, and Company Performance," *Strategic Management Journal,* 12/3 (1991): 219-229.

[7] D. Lovallo and D. Kahneman, "Delusions of success: How Optimism Undermines Executives' Decisions," *Harvard Business Review* (July 2003): 56-63.

[8] Ford (1985), op. cit.; Clapham et al. (1991), op. cit.

[9] R. Buehler, D.W. Griffin, M. Ross, "Exploring the "Planning Fallacy": Why People Underestimate Their Task Completion Times," *Journal of Personality and Social Psychology,* 77 (1994): 366-381.

[10] D. Kahneman and D. Lovallo, "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking," *Management Science,* 39 (1993): 17-31.

[11] Of course, selfishness is a psychological phenomenon, but one that is accepted by neo-classical economists.

[12] D. Kahneman and A. Tversky, "Intuitive Predictions: Biases and Corrective Procedures", in S. Makridakis and S.C. Wheelwright, eds., "Forecasting,", *TIMS, Studies in Management Science,* 12 (1979): 313-327; Kahneman and Lovallo (1993), op. cit.; Lovallo and Kahneman (2003), op. cit.

[13] Lovallo and Kahneman (2003), op. cit.

[14] Buehler et al. (1994), op. cit.

[15] Buehler et al. (1994), op. cit.; R. Buehler, D.W. Griffin and H. MacDonald, "The Role of Motivated Reasoning in Optimistic Time Predictions," *Personality and Social Psychology Bulletin,* 23/3 (1997): 238-247; I.R. Newby-Clark, I. McGregor and M.P Zanna, "Thinking and Caring About Cognitive Inconsistency: When and For Whom Does Attitudinal Ambivalence Feel Uncomfortable?," *Journal of Personality and Social Psychology,* 82 (2002): 157-166.

[16] Flyvbjerg, Holm, and Buhl (2002), op. cit.; Mott MacDonald, *Review of Large Public Procurement in the UK* (HM Treasury, London UK, 2002); National Audit Office, *PFI: Construction Performance*, *Report by the Controller and Auditor General* (HMSO, London, 2003); National Audit Office, *PFI: Construction Performance, Report by the Controller and Auditor General* (HMSO, London, 2005).

[17] U. Malmendier and G.A. Tate, "Who Makes Acquisitions? CEO Overconfidence and Market's Reaction," Stanford Research Paper*,* 2003*,* No. 1798.

---

[18] T. Dune, M.J. Roberts, and L. Samuelson, “Patterns of Firm Entry and Exit in U.S. Manufacturing Industries,” *Rand Journal of Economics*, 19/4 (1988): 495-515; Lovallo and Kahneman (2003), op. cit.

[19] A. Tversky and D. Kahneman, “Judgment Under Uncertainty: Heuristics and Biases,” *Science,* 185 (1974): 1124-1131.

[20] Tversky and Kahneman (1974), op. cit.

[21] K.A. Diekmann, A.E. Tenbrunsel, P.P. Shah, H.A. Schroth and M.H. Bazerman. “The Descriptive and Prescriptive Use of Previous Purchase Price in Negotiations,” *Organizational Behavior and Human Decision Processes*, 66/2 (1996): 179-191.

[22] B. Flyvbjerg, N. Bruzelius, and W. Rothengatter, *Megaprojects and risk: an anatomy of ambition* (New York: Cambridge University Press, 2003).

[23] E.M. Merrow, P.E. Phillips, and C.W. Meyers, *Understanding Cost Growth and Performance Shortfalls in Pioneer Process Plants* (Santa Monica, CA: Rand Corporation, 1981).

[24] M.C. Jensen and W.H. Meckling “Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure,” in M.C. Jensen, *Theory of the Firm: Governance, Residual Claim, and Organizational Forms* (Cambridge, MA: Harvard University Press, 2001).

[25] P. Bolton and D.S. Scharfstein, “Corporate Finance, the Theory of the Firm, and Organizations,” *Journal of Economic Perspectives,* 12/4 (Autumn 1998): 95-114.

[26] B. Flyvbjerg and Cowi, Procedures for Dealing with Optimism Bias in Transport Planning: Guidance Document (London, UK: UK Department for Transport, 2004).

[27] Flyvbjerg (2005), op. cit, p. 50-59.

[28] Flyvbjerg’s personal communication with Kim Utzon.

[29] Flyvbjerg and Cowi (2004), op. cit.

[30] K. Trapenberg Frick, "The Cost of the Technological Sublime: Daring Ingenuity and the New San Francisco-Oakland Bay Bridge," in H. Priemus, B. Flyvbjerg, and B. van Wee, eds., *Decision-Making On Mega-Projects: Cost–benefit Analysis, Planning, and Innovation* (Cheltenham, UK and Northamton, MA, USA: Edward Elgar, 2008).

[31] D.H. Pickrell, *Urban rail transit projects: forecasts versus actual ridership and cost* (Washington, DC: U.S. Department of Transportation, 1990); D.H. Pickrell, “A Desire Named Streetcar: Fantasy and Fact in Rail Transit Promotions and Evaluation,” *Journal of the American Planning Association,* 58/2 (1992): 158-176.

[32] Flyvbjerg and Cowi (2004), op. cit.

[33] A. Altshuler and D. Luberoff, *Mega-Projects: The Changing Politics of Urban Public Investment* (Washington, DC: Brookings Institution, 2003); K.J. Delaney and R. Eckstein, *Public Dollars, Private Stadiums: The Battle Over Building Sports Stadiums* (New Brunswick, NJ: Rutgers University Press, 2003); P.W.G. Morris and G.H. Hough, *The Anatomy of Major Projects: A Study of the Reality of Project Management* (New York: John Wiley and Sons, 1987).

[34] Ibid.

[35] HM Treasury, *The Green Book: Appraisal and Evaluation in Central Government* (London, UK: Treasury Guidance TSO, 2003); Flyvbjerg and Cowi (2004), op. cit.

[36] B. Flyvbjerg, M.S. Holm and S. Buhl, "What Causes Cost Overrun in Transport Infrastructure Projects?,"

Transport Reviews, 24/1 (2004): 3-18.

[37] The problem of accountability in public and private organizations has been extensively discussed in M. Bovens, *The Quest for Responsibility: Accountability and Citizenship in Complex Organisations* (Cambridge, U.K.: Cambridge University Press, 1998) and M. Bovens, “Analysing and Assessing Accountability: A Conceptual Framework,” *European Law Journal:* 13/4 (2007): 447-468.

[38] Kahneman and Lovallo (1993), op. cit.

[39] A. Cooper, C. Woo and W. Dunkelberg, “Entrepreneurs’ Perceived Chances of Success,” *Journal of Business Venturing,* 3 (1988): 97-108.

[40] J. Farrell, “Cheap Talk, Coordination and Entry,” *Rand Journal of Economics* 18/1 (1987): 34-39

[41] Flyvbjerg et al. (2002), op. cit.

[42] Flyvbjerg et al. (2005), op. cit.

[43] W. Chen, "Analysis of Rail Transit Project Selection Bias With an Incentive Approach," *Planning Theory*, vol. 6, no. 1, 2007, pp. 69-94.

[44] Flyvbjerg et al. (2005), op. cit., p. 139-40.

[45] Flyvbjerg et al. (2005), op. cit.

[46] In this section, “incentive” is used in a very broad sense, including changes in governance structures and control mechanisms. We thank an anonymous reviewer for pointing this out.

[47] Flyvbjerg (2006), op. cit., 5-15.

[48] C. Caringal, D. Lovallo and C. Camerer, “Similarity Based Forecasting,” AGSM Working Paper, 2007.

[49] B. Flyvbjerg, "Policy and Planning for Large-Infrastructure Projects: Problems, Causes, Cures," *Environment and Planning B: Planning and Design*, vol. 34, 2007, pp. 578-597.

[50] Flyvbjerg, Bruzelius, and Rothengatter (2003), op. cit.

[51] Department for Transport, “Consultation on Changes to the Policy on Funding Major Projects: Local Authority Major Schemes,” 2006a, London.

[52] Ibid, par 5.

[53] Ibid, par. 38-39.

[54] Department for Transport, “Transport Analysis Guidance (TAG), Unit 3.5.9, The Estimation and Treatment of Scheme Costs,” 2006b, London, par. 3.3.7.

[55] Flyvbjerg, Bruzelius, and Rothengatter (2003), op. cit., p. 120-123.

[56] In the US, after the Sarbanes-Oxley Act was passed in 2002 in response to corporate scandals with Enron, Tyco, WorldCom, etc., CEOs and CFOs of publicly traded companies have become criminally liable for misleading shareholders about financial projections, including cost, revenue, and risk estimates.

[57] On the contrary, the “winner’s curse” reflects the fact that the winning bid exceeds the real value of the prize.

[58] In a Public Private Partnership (PPP), a single contractual arrangement combine a series of separate arrangements, such as the design, building, financing, ownership, maintenance, etc. DBFOM is only an example of it. Build-Own-Operate and Transfer (BOOT) and Build-Own and Operate (BOO) indefinitely are other cases. These arrangements are often contrasted with traditional procurement methods, such as Design and Build (DB), where a private entity engages only in the design and building of the infrastructure.

Mott Macdonald (2002), op. cit. discusses performance of traditional procurement and PPP in 50 large UK infrastructure projects.

[59] D. Kahneman, “New Challenges to the Rationality Assumption,” *Journal of Institutional and Theoretical Economics*, 150 (1994): 18-36; Kahneman and Tversky (1979), op. cit.; Lovallo and Kahneman (2003), op. cit.; Flyvbjerg (2006), op. cit.

[60] Kahneman and Tversky (1979), op. cit.

[61] Kahneman and Lovallo (1993), op. cit.

[62] In judging the likelihood or frequency of an event, decision makers tend to rely on information that is easy to recall. On the availability heuristic, see A. Tversky and D. Kahneman, “Availability: A Heuristic for Judging Frequency and Probability,” *Cognitive Psychology,* 5 (1973), 207-232.

[63] This discussion is based on Kahneman and Tversky (1979), op. cit.

[64] B. Flyvbjerg and Cowi, Procedures for Dealing with Optimism Bias in Transport Planning: Guidance Document (London, UK: UK Department for Transport, 2004).

[65] Ove Arup & Partners Scotland, *Scottish Parliament, Edinburgh Tram Line 2 Review of Business Case* (Scotland: West Lothian, 2004), pp. 15-16.

[66] Caringal et al. (2007), op. cit.

[67] M. Wachs, “Ethics and Advocacy in Forecasting for Public Policy,” *Business and Professional Ethics Journal,* 9/1-2 (1990): 141-157.

[68] L. Merewitz, “Cost Overruns in Public Works,” in W. Niskanen, A.C. Hansen, R.H. Havemann, R. Turkey, R. Zeckhauser, eds., *Benefit cost and policy analysis*: 277-295 (Chicago IL: Aldine, 1973), p. 280.

[69] However, such explanations depict misrepresentation as deliberate, as well as politically and economically rational. If we now define a lie in the conventional fashion as making a statement intended to deceive others, we see that deliberate misrepresentation of costs and benefits is lying. If such lying is done for the public good (e.g., to save taxpayers’ money), political theory would classify it in that special category of lying called the “noble lie,” the lie motivated by altruism. According to expert on lying, Sissela Bok, this is the “most dangerous body of deceit of all”. See S. Bok, *Lying: Moral Choice in Public and Private Life* (New York: Vintage, 1979), p. 14 and p. 175; L. Cliffe, M. Ramsey and D. Bartlett, *The Politics of Lying: Implications for Democracy* (London: Macmillan, 2000), p. 3.

[70] H. Priemus, B. Flyvbjerg, and B. van Wee, eds. (2003), op. cit.

[71] R. Anguera, "The Channel Tunnel: An Ex Post Economic Evaluation," *Transportation Research Part A*, 40 (2006): 291–315.

Figure 1: Pioneer process plants cost forecasts accuracy (Merrow, Phillips and Meyers, 1981)

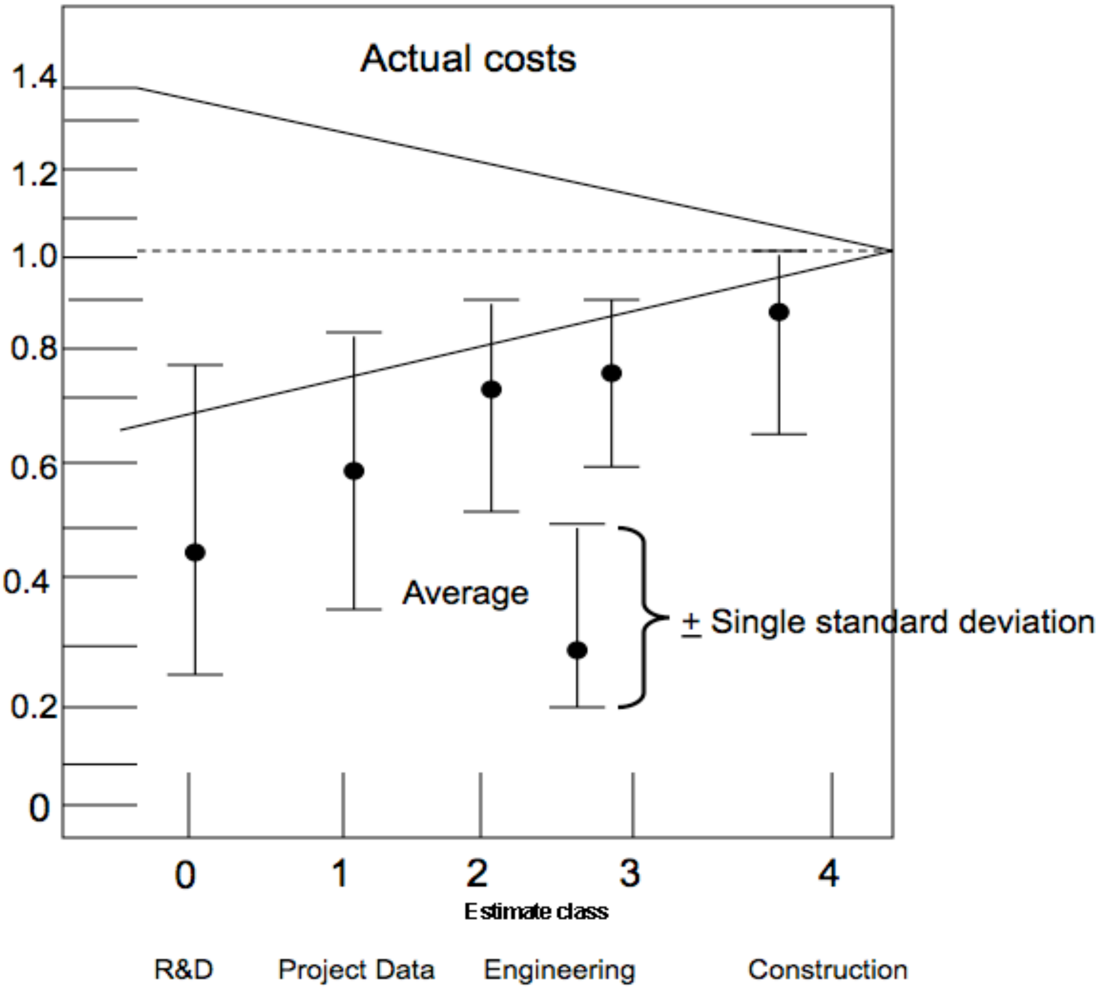


Figure 2: Multi-tier principal-agent relationships

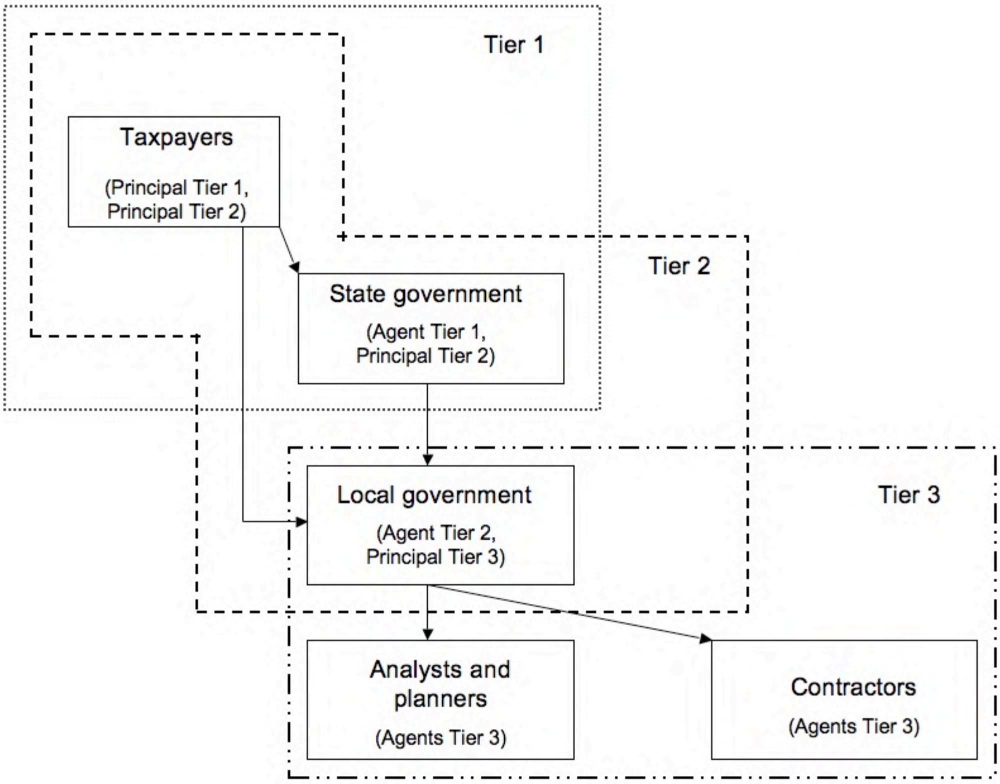


Figure 3: Incentives and learning in large capital investments

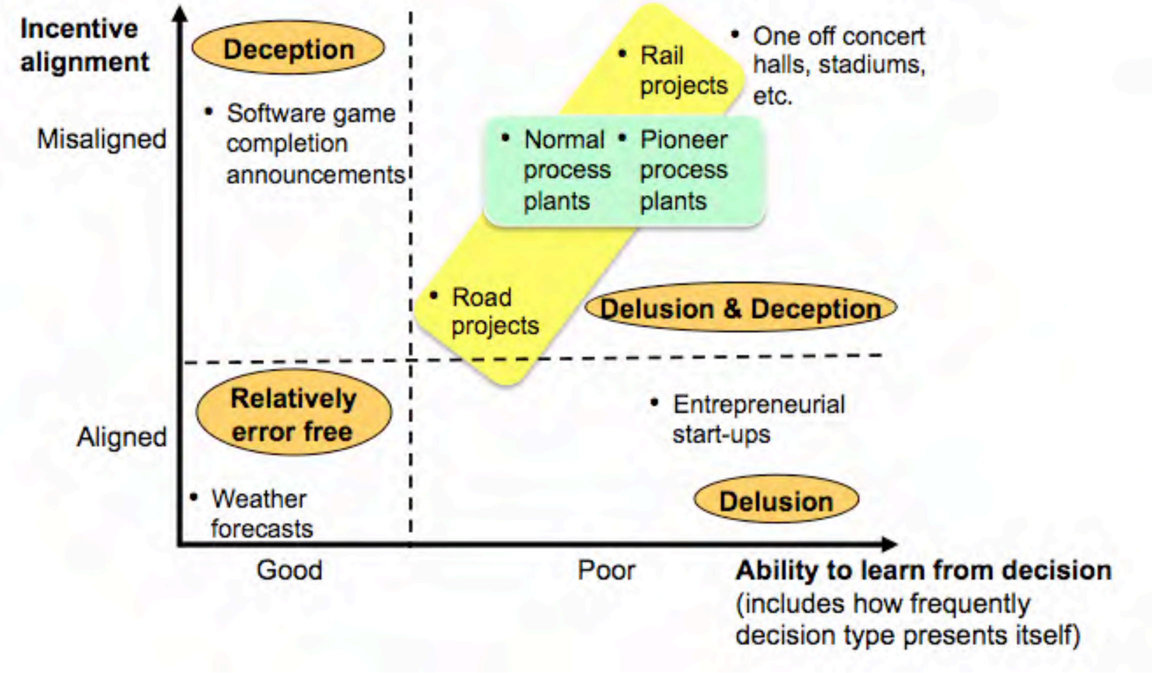


Table 1: Avoiding strategic deception

Actors	Causes of deception	Prescriptive advice
1. Proposing and approving institutions (i.e. agents of the taxpayers)	Proposing institution strategically misrepresents costs, timeframe, risks and benefits to obtain funding	Proposing and approving institutions should share financial responsibility (e.g. minimum local contribution, local contribution for cost increases, identification of who "owns" the risks) Private financers should participate, without a sovereign guarantee, for a least one third of the total capital needs
2. Planners, bidders and contractors (i.e. agents of the proposing institution)	Planners strategically misrepresent costs, timeframe and benefits (to please the proposing institution)	Financial and non-financial rewards for planners who proposed realistic estimates Strict forecasts audit Criminal penalties for purposely misleading forecasts
	Bidders propose artificially low bids because of planned compensation through expected scope increases	Place financial responsibility with bidders
	Contractors overprice scope increases	Place financial responsibility with contractors for delays and scope increases

TAB 2

Energy Policy ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Energy Policy

journal homepage: www.elsevier.com/locate/enpol

Should we build more large dams? The actual costs of hydropower megaproject development ☆

Atif Ansar [a,b,*], Bent Flyvbjerg [b], Alexander Budzier [b], Daniel Lunn [c]

[a] Blavatnik School of Government, University of Oxford, Oxford OX1 4JJ, UK
[b] Saïd Business School, University of Oxford, Oxford OX1 1HP, UK
[c] Department of Statistics, University of Oxford, Oxford OX1 3GT, UK

HIGHLIGHTS

- We investigate *ex post* outcomes of schedule and cost estimates of hydropower dams.
- We use the "outside view" based on Kahneman and Tversky's research in psychology.
- Estimates are systematically and severely biased below actual values.
- Projects that take longer have greater cost overruns; bigger projects take longer.
- Uplift required to de-bias systematic cost underestimation for large dams is +99%.

ARTICLE INFO

Article history:
Received 25 February 2013
Accepted 29 October 2013

Keywords:
Large hydropower dams
Schedule and cost estimates
Cost–benefit forecasting
Reference class forecasting
Outside view

ABSTRACT

A brisk building boom of hydropower mega-dams is underway from China to Brazil. Whether benefits of new dams will outweigh costs remains unresolved despite contentious debates. We investigate this question with the "outside view" or "reference class forecasting" based on literature on decision-making under uncertainty in psychology. We find overwhelming evidence that budgets are systematically biased below actual costs of large hydropower dams—excluding inflation, substantial debt servicing, environmental, and social costs. Using the largest and most reliable reference data of its kind and multilevel statistical techniques applied to large dams for the first time, we were successful in fitting parsimonious models to predict cost and schedule overruns. The outside view suggests that in most countries large hydropower dams will be too costly in absolute terms and take too long to build to deliver a positive risk-adjusted return unless suitable risk management measures outlined in this paper can be affordably provided. Policymakers, particularly in developing countries, are advised to prefer agile energy alternatives that can be built over shorter time horizons to energy megaprojects.

1. Large hydropower dam controversy

The 21st Century faces significant energy challenges on a global scale. Population and economic growth underpin increasing demand for energy from electricity to transport fuels. Social objectives of poverty alleviation, adaptation and mitigation of climate change, and energy security present policy makers and business leaders with difficult decisions and critical trade-offs in implementing sound energy policies. Demand for electricity is, for example, slated to almost double between 2010 and 2035 requiring global electricity capacity to increase from 5.2 terawatt (TW) to 9.3 TW over the same period (IEA, 2011). Currently, the de facto strategic response to these big energy challenges is "big solutions" such as large hydropower dams. Are such big solutions in general and large hydropower dams in particular the most effective strategy, on a risk-adjusted basis, to resolve global energy challenges? Might more numerous small interventions be more prudent from the perspective of risk management and maximizing net present value even when they entail somewhat higher per unit cost of production?

Proponents of large dams envisage multiple benefits. A big step-up in hydropower capacity along with a long and varied list of corollary benefits: reducing fossil fuel consumption, flood control, irrigation, urban water supply, inland water transport, technological progress, and job creation (Billington and Jackson, 2006; ICOLD, 2010). Inspired by the promise of prosperity, there is a robust

---

☆ This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-No Derivative Works License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

* Corresponding author. Tel.: +44 186 528 8795; fax: +44 186 528 8805.
E-mail addresses: atif.ansar@bsg.ox.ac.uk, atifansar@gmail.com (A. Ansar).

0301-4215/$ - see front matter © 2014 The Authors. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.enpol.2013.10.069

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

pipeline of new mega-dams being developed globally after a two-decade lull. The Belo Monte dam in Brazil, the Diamer-Bhasha in Pakistan, Jinsha river dams in China, Myitsone dam in Myanmar, or the Gilgel Gibe III dam in Ethiopia, all in various stages of development, are unprecedented in scale.

Large dams are, however, controversial because they exert substantial financial costs (World Bank, 1996; World Commission on Dams, 2000). Beyond the financial calculus, large dams have profound environmental (McCully, 2001; Scudder, 2005; Stone, 2011), ecological (Nilsson et al., 2005; Ziv et al., 2012), and social (Bakker, 1999; Duflo and Pande, 2007; Richter et al., 2010; Sovacool and Bulan, 2011) impacts. Stone (2011, p. 817) reports in *Science* that the Three Gorges dam in China is an “environmental bane” that will cost over USD 26.45 billion over the next 10 years in environmental “mitigation efforts”. Despite their outsized financial and environmental costs, the purported benefits of large hydropower dams prove uncertain. For example, the World Commission of Dams (2000, p. 30) reported that for large hydropower dams “average [hydropower] generation in the first year of commercial operation is 80% of the targeted value”—a trend of which the recently completed Bakun hydroelectric project in Borneo is an alarming example (Sovacool and Bulan, 2011). Similarly, Duflo and Pande (2007) find adverse distributional impacts of large irrigation dams in India. Winners downstream come with losers upstream yielding a more modest, if any, net economic benefit.

The scale of contemporary large dams is so vast that even for a large economy such as China's the negative economic ramifications “could likely hinder the economic viability of the country as a whole” if the risks inherent to these projects are not well managed (Salazar, 2000). Similarly, Merrow et al. (1988, pp. 2–3) warn that “such enormous sums of money ride on the success of megaprojects [such as large dams] that company balance sheets and even government balance-of-payments accounts can be affected for years by the outcomes”. Such warnings are not idle alarmism. There is mounting evidence in civil society, academic research, and institutional accounts that large dams have strikingly poor performance records in terms of economy, social and environmental impact, and public support (McCully, 2001; Scudder, 2005; Singh, 2002; Sovacool and Bulan, 2011; WCD, 2000). There are acrimonious, and as yet inconclusive, debates in scientific literature and civil society about whether large dams are a boon or a curse. Should we build more large hydropower dams? How confident can planners be that a large bet on a large dam will pay-off handsomely?

We investigate these questions with the “outside view” or “reference class forecasting” based on the literature on decision-making under uncertainty that won Princeton psychologist Daniel Kahneman the Nobel Prize in economics in 2002 (Kahneman and Tversky, 1979a, 1979b; Kahneman, 1994) extended and applied by Bent Flyvbjerg and colleagues to infrastructure projects (Flyvbjerg et al., 2003; Flyvbjerg, 2009). We present statistical and comparative evidence from the largest reference class to-date of actual costs of large hydropower dam projects (hereafter large dams unless stated otherwise). We find that even before accounting for negative impacts on human society and environment, the actual construction costs of large dams are too high to yield a positive return. Large dams also take inordinately long periods of time to build, making them ineffective in resolving urgent energy crises. Our evidence pertains primarily to large dams and the results cannot be applied either to smaller dams or other large energy solutions such as nuclear power without first building a separate “reference class” for other types of power generation technologies. Our findings, however, point towards the generalizable policy proposition that policymakers should prefer energy alternatives that require less upfront outlays and that can be built very quickly.

There is no doubt that harnessing and managing the power of water is critical for economies but large dams are not the way to do so unless suitable risk management measures outlined in this paper can be affordably provided. Building on literature in decision making under uncertainty in management, psychology, and planning research, this paper further provides public agencies (e.g. national planning and finance ministries, power and water authorities), private entrepreneurs, investors, and civil society a framework to test the reliability of *ex ante* estimates for construction costs and schedules of power generation alternatives. An impartial and rigorous application of the reference class forecasting methods proposed here can improve the selection and implementation of new investments.

2. Delusion and deception in large hydropower dam planning?

Our approach to address the debates about whether or not to build dams is to incorporate an evidence-based perspective that reflects how decisions among alternative options are actually made and on what basis. Theoretical and empirical literature on decision-making under uncertainty proposes two explanations—psychological delusion and political deception—that suggest decision-makers' forecasts, and hence ex ante judgments, are often adversely biased (Tversky and Kahneman, 1974; Kahneman and Lovallo, 1993; Flyvbjerg, 2003; Lovallo and Kahneman, 2003; Kahneman, 2011).

First, experts (e.g., statisticians, engineers, or economists) and laypersons are systematically and predictably too optimistic about the time, costs, and benefits of a decision. This “planning fallacy” (Kahneman and Tversky, 1979b; Buehler et al., 1994) stems from actors taking an “inside view” focusing on the constituents of the specific planned action rather than on the outcomes of similar actions already completed (Kahneman and Lovallo, 1993). Thus, for example, the estimated costs put forward by cities competing to hold the Olympic Games have consistently been underestimated yet every four years these errors are repeated. Biases, such as overconfidence or overreliance on heuristics (rules-of-thumb), underpin these errors.

Second, optimistic judgments are often exacerbated by deception, i.e. strategic misrepresentation by project promoters (Wachs, 1989; Pickrell, 1992; Flyvbjerg et al., 2002, 2005, 2009). Recent literature on infrastructure delivery finds strong evidence that misplaced political incentives and agency problems lead to flawed decision-making (see Flyvbjerg et al., 2009). Flyvbjerg et al. (2009, p. 180) further discuss that delusion and deception are complementary rather than alternative explanations for why megaprojects typically face adverse outcomes. It is, however, “difficult to disentangle” delusion from deception in practice. Using quasi-experimental evidence from China, Ansar et al. (2013) suggest that while better incentive alignment can help to lower the frequency and, to a lesser extent, the magnitude of biases, it does not entirely cure biases.

Be it delusion or deception, is decision-making in large hydropower dams systematically biased by errors in cost, schedule, and benefit forecasts? What is the risk that costs might outweigh benefits for a proposed dam? While the future is unknowable, uncertain outcomes of large investments can still be empirically investigated using “reference class forecasting” (RCF) or the “outside view” techniques (Kahneman and Lovallo, 1993; Flyvbjerg, 2006, 2008). To take an outside view on the outcome of an action (or event) is to place it in the statistical distribution of the outcomes of comparable, already-concluded, actions (or events). The outside view has three advantages: First, it is evidence-based and requires no restrictive assumptions. Second, it helps to test and fit models to explain why the outcomes of a reference class of

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

past actions follow the observed distribution. Third, it allows to predict the uncertain outcomes of a planned action by comparing it with the distributional information of the relevant reference class. The theoretical foundations of the outside view were first described by Kahneman and Tversky (1979b) and later by Kahneman and Lovallo (1993) and Lovallo and Kahneman (2003) as means to detect and cure biases in human judgment. The methodology and data needed for employing the outside view, or reference class forecasting, in practice were developed by Flyvbjerg (2006, 2008) in collaboration with the Danish consulting firm COWI (see Flyvbjerg and COWI, 2004).

2.1. *Three steps to the outside view*

The outside view, applied to large dams for the first time here, involves three steps: (i) identify a reference class; (ii) establish an empirical distribution for the selected reference class of the parameter that is being forecasted; (iii) compare the specific case with the reference class distribution. We take a further innovatory step of fitting multivariate multilevel models to the reference data to predict future outcomes. Our technique is an important improvement in the methodology of the outside view that can be generalized and applied to other large-scale and long-term decisions under uncertainty. With de-biased forecasts managers can make empirically and statistically grounded, rather than optimistic, judgments (Dawes et al., 1989; Buehler et al., 1994; Gilovich et al., 2002).

The outside view—as implemented by Flyvbjerg (2006, 2008)—is not without limitations (see Sovacool and Cooper, 2013 for a discussion specifically about energy megaprojects). For example, RCF focuses on generic risk inherent in a reference class rather than specific project-level risk. We rectify against this limitation by fitting regression models in addition to using traditional RCF methods in the result section below. Sovacool and Cooper (2013, p. 63) further suggest that RCF may not provide sufficiently accurate indication of the risks of rare megaprojects the likes of which have never been built before. Such “out of the sample” problems are well noted in probability theory. They do not, however, deny the fundamental usefulness of RCF. If anything our results err towards conservative estimates of actual cost overruns and risks experienced by large dams.

2.2. *Measures and data*

Following literature on the planning fallacy (Sovacool and Cooper, 2013), the parameters central to our investigation and multilevel regression analysis is the inaccuracy between managers' forecasts and actual outcomes related to construction costs, or the cost overrun, and implementation schedule, or schedule slippage. Following convention, cost overrun is the actual outturn costs expressed as a ratio of estimated costs[1]; cost overruns can also be thought as the underestimation of actual costs (Bacon and Besant-Jones, 1998; Flyvbjerg et al., 2002). Schedule slippage, called schedule overrun, is the ratio of the actual project implementation duration to the estimated project implementation. The start of the implementation period is taken to be the date of project approval by the main financiers and the key decision makers, and the end is the date of full commercial operation.

Inaccuracies between actual outcomes versus planned forecasts are useful proxies for the underlying risk factors that led to the inaccuracies. For example, cost overruns reduce the attractiveness of an investment and if they become large the fundamental economic viability becomes questionable. Bacon and Besant-Jones (1998, p. 317) offer an astute summary:

> The economic impact of a construction cost overrun is the possible loss of the economic justification for the project. A cost overrun can also be critical to policies for pricing electricity on the basis of economic costs, because such overruns would lead to underpricing. The financial impact of a cost overrun is the strain on the power utility and on national financing capacity in terms of foreign borrowings and domestic credit.

Similarly, schedule slippages delay much needed benefits, expose projects to risks such as an increase in finance charges, or creeping inflation, which may all require upward revision in nominal electricity tariffs. Financial costs and implementation schedules, because of their tangibility, are also good proxies for non-pecuniary impacts such as those on the environment or on the society. Projects with a poor cost and schedule performance are also likely to have a poor environmental and social track record. A greater magnitude of cost and schedule overruns is thus a robust indicator of project failure (Flyvbjerg, 2003).

In taking the outside view on the cost and schedule under/overruns, our first step was to establish a valid and reliable reference class of previously built hydropower dams as discussed above. The suggested practice is that a reference class ought to be broad and large enough to be statistically meaningful but narrow enough to be comparable (Kahneman and Tversky, 1979b; Kahneman and Lovallo, 1993; Flyvbjerg, 2006). International standard defines dams with a wall height > 15 m as large. The total global population of large dams with a wall height > 15 m is 45,000. There are 300 dams in the world of monumental scale; these “major dams” meet one of three criteria on height (> 150 m), dam volume (> 15 million m^3), or reservoir storage (> 25 km^3) (Nilsson et al., 2005).

From this population of large dams, our reference class drew a representative sample of 245 large dams (including 26 major dams) built between 1934 and 2007 on five continents in 65 different countries—the largest and most reliable data set of its kind. The portfolio is worth USD 353 billion in 2010 prices. All large dams for which valid and reliable cost and schedule data could be found were included in the sample. Of the 245 large dams, 186 were hydropower projects (including 25 major dams) and the remaining 59 were irrigation, flood control, or water supply dams. While we are primarily interested in the performance of large dam projects with a hydropower component, we also included non-hydropower dam projects in our reference class to test whether project types significantly differ in cost and schedule overruns or not. Fig. 1 presents an overview of the sample by regional location, wall height, project type, vintage, and actual project cost.

The empirical strategy of this paper relied on documentary evidence on estimated versus actual costs of dams. Primary documents were collected from *ex ante* planning and *ex post* evaluation documents of the:

1. Asian Development Bank;
2. World Bank, also see World Bank (1996) and Bacon and Besant-Jones (1998);
3. World Commission of Dams (WCD), also see WCD (2000)[2];
4. U.S. Corps of Engineers;
5. Tennessee Valley Authority;

---

[1] Cost overruns can also be expressed as the actual outturn costs minus estimated costs in percent of estimated costs.

[2] Note that the World Bank, Asian Development Bank, and the WCD typically report cost data in nominal USD. We, however, converted these data, adapting methods from World Bank (1996: 85), into constant local currencies.

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

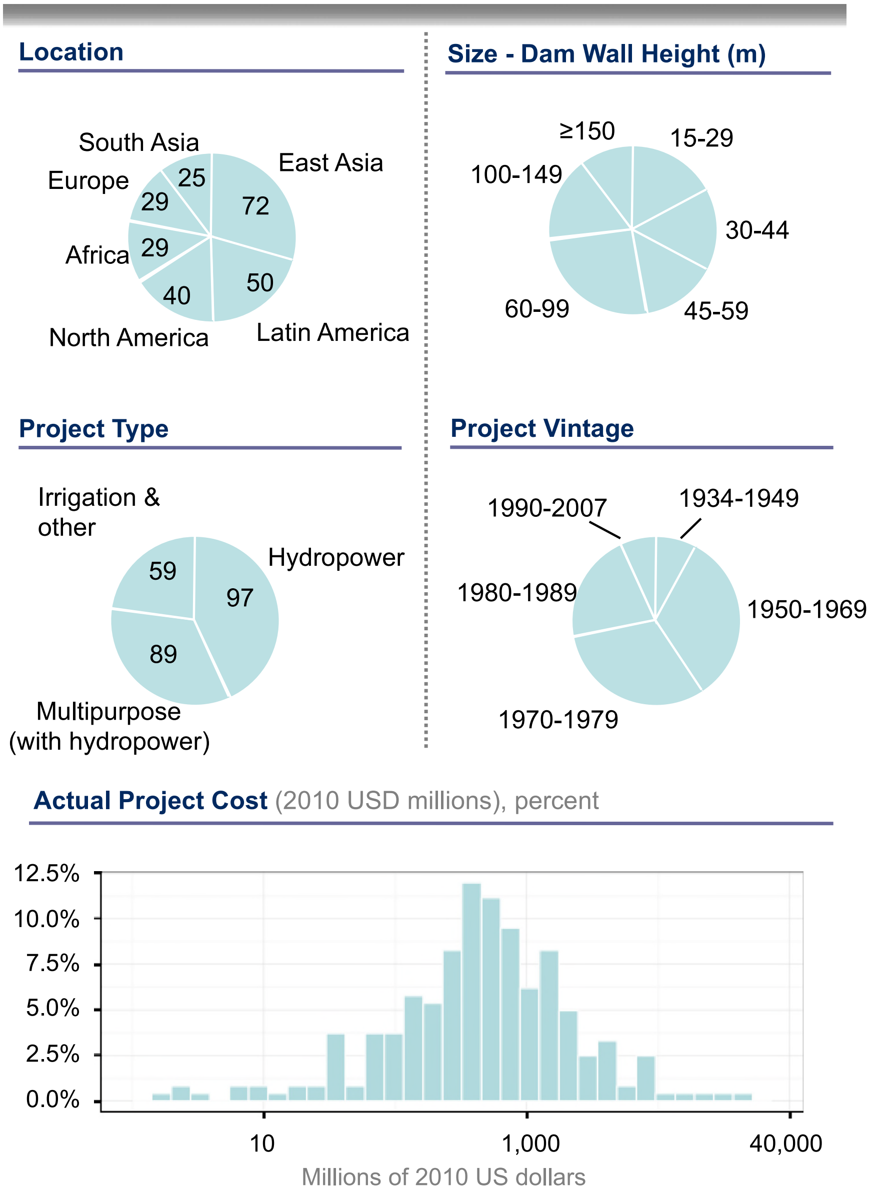


Fig. 1. Sample distribution of 245 large dams (1934–2007), across five continents, worth USD 353B (2010 prices).

6. U.S. Bureau of Reclamation, also see Hufschmidt and Gerin (1970)[3] and Merewitz (1973) on the U.S. water-resource construction agencies.

The procedures applied to the cost and schedule data here are consistent with the gold standard applied in the field—more detailed methodological considerations can be found in Flyvbjerg et al. (2002), Federal Transit Administration (2003), Pickrell (1989, 1992), World Bank (1996) and Bacon and Besant-Jones (1998) with which our data are consistent. All costs are total project costs comprising the following elements: right-of-way acquisition and resettlement; design engineering and project management services; construction of all civil works and facilities; equipment purchases. Actual outturn costs are defined as real, accounted construction costs determined at the time of project completion. Estimated costs are defined as budgeted, or forecasted, construction costs at the time of decision to build. The year of the date of the decision to build a project is the base year of prices in which all estimated and actual constant costs have been expressed in real (i.e. with the effects of inflation removed) local currency terms of the country in which the project is located. We exclude from our calculations debt payments, any *ex post* environmental remedial works, and opportunity cost of submerging land to form reservoirs. This makes comparison of estimated and actual costs of a specific project a like-for-like comparison.

2.3. Analyses

We investigated the magnitude and frequency of cost and schedule forecast (in)accuracies with a combination of simple

---

[3] Hufschmidt and Gerin (1970) report data on over 100 dams built in the United States between 1933 and 1967. The salient results of the study were that in nominal USD terms dams built by TVA suffered a 22% cost overrun; U.S. Corps of Engineers overrun was 124% for projects built or building prior to 1951, and 36% for projects completed between 1951 and 1964; while U.S. Bureau of Reclamation overrun was 177 per cent for projects built or building prior to 1955 and 72 per cent for all projects built or building in 1960 (Hufschmidt and Gerin, 1970: 277). Despite its large sample, Hufschmidt and Gerin (1970) do not report data broken down project-by-project. The validity and reliability of these data could not thus be established and were consequently excluded.

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

statistical (parametric and non-parametric) tests and by fitting more sophisticated multilevel regression models sometimes termed Hierarchical Linear Models (HLM).

Multilevel or hierarchically structured data are the norm in the social, medical, or biological sciences. Rasbash et al. (2009, p. 1) explain: “For example, school education provides a clear case of a system in which individuals are subject to the influences of grouping. Pupils or students learn in classes; classes are taught within schools; and schools may be administered within local authorities or school boards. The units in such a system lie at four different levels of a hierarchy. A typical multilevel model of this system would assign pupils to level 1, classes to level 2, schools to level 3 and authorities or boards to level 4. Units at one level are recognized as being grouped, or nested, within units at the next higher level. Such a hierarchy is often described in terms of clusters of level 1 units within each level 2 unit, etc. and the term clustered population is used.” Important for a hierarchical linear model is that the dependent variable is at the lowest level of the nested structure. Multilevel models are necessary for research designs where data for observations are organized at more than one level (i.e., nested data) (Gelman and Hill, 2007). Failing to use multilevel models in such instances would result in spurious results (Rasbash et al., 2009).

With respect to our data on dams, projects are nested in the countries of their domicile. Like test scores of pupils from the same school tend exhibit within-school correlation, similarly outcomes of dam projects may exhibit within-country correlation that needs to be properly modeled using a multilevel model. We took this into account by modeling country as a first level random effect in a mixed effects multilevel model. The models were made parsimonious by using stepwise variable selection.

3. Results and interpretation

Our second step was to establish an empirical distribution for the cost forecast errors of large dams. We collected data on 36 possible explanatory variables, listed in Table 1, for the 245 large dams in our reference class.

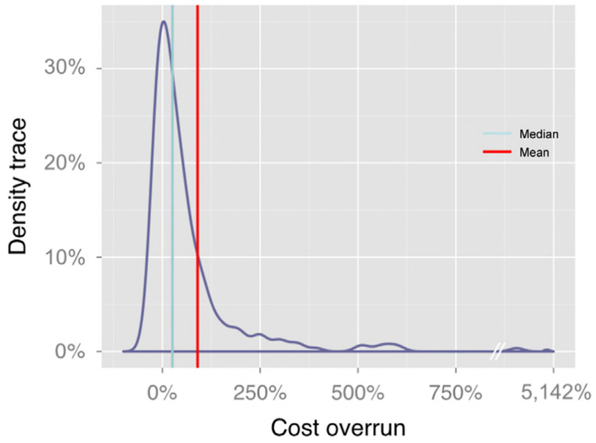


Fig. 2. Density trace of actual/estimated cost (i.e. costs overruns) in constant local currency terms with the median and mean ($N=245$).

Table 1
Variables and characteristics used in multilevel regressions on construction cost overrun and schedule slippage.

Project-specific variables
Project features
Hydropower or non-hydropower large dam project (dummy variable)
New power station or station extension (dummy variable)
Size
Generator unit capacity (MW)
Total project generation capacity (MW)
Dam height for new hydropower station (meters)
Hydraulic head for new hydropower station (meters)[a]
Reservoir area created by project (hectares)[a]
Length of tunnels (kilometers)[a]
Cost
Estimated project cost (constant local currency converted to 2010 USD MM)
Actual project cost (constant local currency converted to 2010 USD MM)
Cumulative inflation contingency (percentage)
Time
Year of final decision to build
Estimated implementation schedule (months)
Year of start of full commercial operation
Actual implementation schedule (months)
Procurement
Estimated project foreign exchange costs as a proportion of estimated total project costs (percentage)
Competitiveness of procurement process, international competitive bidding amount as a proportion of estimated total project costs (percentage)[*]
Main contractor is from the host country (dummy variable)
Country variables
Country (second level to control for within country correlation)
Political regime of host country is a democracy (dummy variable)
GDP of host country (current USD)
Per capita income of host country in year of loan approval (constant USD)
Average actual cost growth rate in host country over the implementation period—the GDP deflator (percentage)
MUV Index of actual average cost growth rate for imported project components between year of loan approval and year of project completion
Long-term inflation rate of the host country (percentage)
Actual average exchange rate depreciation or appreciation between year of formal-decision-to-build and year of full commercial operation (percentage)
South Asian projects (dummy variable)
North American projects (dummy variable)

[a] Denotes variables with a large number of missing values not used for regression analysis.

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

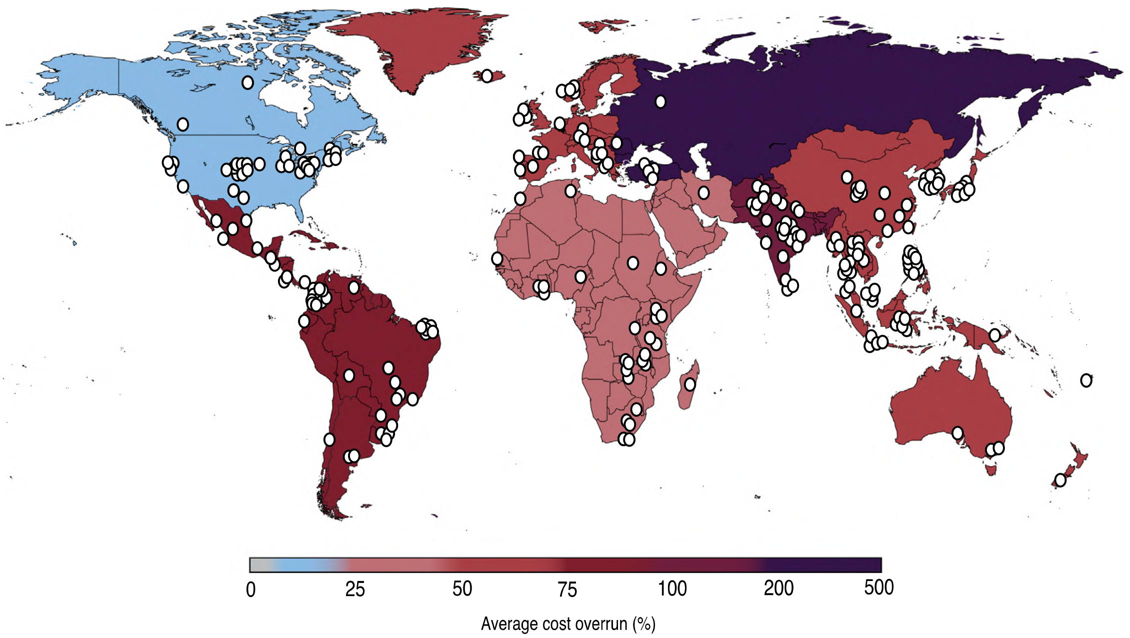


Fig. 3. Location of large dams in the sample and cost overruns by geography.

3.1. Preliminary statistical analysis of cost performance

With respect to cost overruns, we make the following observations:

1. Three out of every four large dams suffered a cost overrun in constant local currency terms.
2. Actual costs were on average 96% higher than estimated costs; the median was 27% (*IQR* 86%). The evidence is overwhelming that costs are systematically biased towards underestimation (Mann–Whitney–Wilcoxon $U=29{,}646$, $p<0.01$); the magnitude of cost underestimation (i.e. cost overrun) is larger than the error of cost overestimation ($p<0.01$). The skew is towards adverse outcomes (i.e. going over budget).
3. Graphing the dams' cost overruns reveals a fat tail as shown in Fig. 2; the actual costs more than double for 2 out of every 10 large dams and more than triple for 1 out of every 10 dams. The fat tail suggests that planners have difficulty in computing probabilities of events that happen far into the future (Taleb, [2007] 2010, p. 284).
4. Large dams built in every region of the world suffer systematic cost overruns. The mean forecasting error is significantly above zero for every region. Fig. 3 shows the geographical spread and cost overruns of large dams in our reference class. Large dams built in North America ($N=40$) have considerably lower cost overrun ($M=11\%$) than large dams built elsewhere ($M=104\%$). Although after controlling for other covariates such as project scale in a multilevel model, reported below, the differences among regions are not significant. We noted, three out of four dams in our reference class had a North American firm advising on the engineering and economic forecasts. Consistent with anchoring theories in psychology, we conjecture that an over-reliance on the North American experience with large dams may bias cost estimates downwards in rest of the world. Experts may be “anchoring” their forecasts in familiar cases from North America and applying insufficient “adjustments” (Flyvbjerg et al., 2009; Tversky and Kahneman, 1974), for example to adequately reflect the risk of a local currency depreciation or the quality of local project management teams. Instead of optimistically hoping to replicate the North American cost performance, policymakers elsewhere ought to consider the global distributional information about costs of large dams.
5. The typical forecasted benefit-to-cost ratio was 1.4. In other words, planners expected the net present benefits to exceed the net present costs by about 40%. Nearly half the dams suffered a cost overrun ratio of 1.4 or greater breaching this threshold after which the asset can be considered stranded—i.e. its upfront sunk costs are unlikely to be recovered. This is assuming, of course, that the benefits did not also fall short of targets, even though there is strong evidence that actual benefits of dams are also likely to fall short of targets (WCD, 2000; McCully, 2001; Scudder, 2005).[4]
6. We tested whether forecasting errors differ by project type (e.g., hydropower, irrigation, or multipurpose dam) or wall type (earthfill, rockfill, concrete arch, etc.). Pairwise comparisons of percentage mean cost overrun and standard deviations as well as non-parametric Mann–Whitney tests for each of the parameters show no statistically significant differences. We conclude that irrespective of project or wall type, the probability distribution from our broader reference class of 245 dams applies as in Fig. 2.
7. We analyzed whether cost estimates have become more accurate over time. Statistical analysis suggests that irrespective of the year or decade in which a dam is built there are no significant differences in forecasting errors ($F=0.57$, $p=0.78$). Similarly, there is no linear trend indicating improvement or deterioration of forecasting errors ($F=0.54$, $p=0.46$) as also suggested in Fig. 4. There is little learning from past mistakes. By the same token, forecasts of costs of large dams today are likely to be as wrong as they were between 1934 and 2007.

We also explored the absolute costs of large hydropower dams ($N=186$). A large hydropower dam on average costs 1800 million in 2010 USD with an average installed capacity of 630 MW. One MW installed capacity on average costs 2.8 million in 2010 USD.

---

[4] A more comprehensive inquiry into planned versus actual benefits of dams is postponed until a future occasion but data available on 84 of the 186 large hydroelectric dam projects thus far suggests that they suffer a mean benefits shortfall of 11%.

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

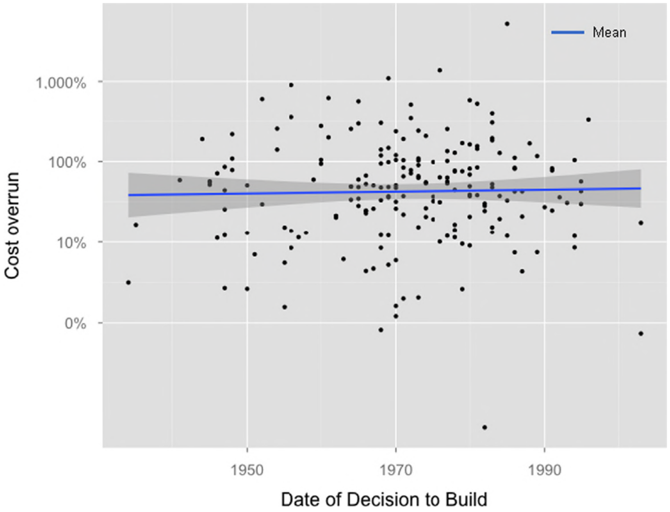


Fig. 4. Inaccuracy of cost estimates (local currencies, constant prices) for large dams over time ($N=245$), 1934–2007.

A preliminary univariate analysis, which makes no attempts to take into account any covariates, shows that increase in the scale of a dam, e.g., measured as height of the dam wall, increases the absolute investment required exponentially, e.g. a 100 m high dam wall is four times more costly than a 50 m wall ($R^2=0.27$, $F=92.5$, $p<0.01$). An even stronger relationship can be seen between installed capacity MW and actual costs ($R^2=0.70$, $F=461.1$, $p<0.01$).

Furthermore, the rate of cost overrun outliers increases with increase in dam size either measured in installed hydropower generation ($r=0.24$, $p=0.01$) or wall height ($r=0.13$, $p=0.05$). Since there is a significant correlation between dam height and hydropower installed capacity ($r=0.47$, $p<0.01$), evidence suggests that larger scale in general is prone to outlying cost overruns. We further investigate the effects of scale on cost overruns by fitting multilevel models (Models 1 and 2) reported below.

3.2. Preliminary statistical analysis of schedule performance

Not only are large dams costly and prone to systematic and severe budget overruns, they also take a long time to build. Large dams on average take 8.6 years. With respect to schedule slippage, we make the following observations:

8. Eight out of every 10 large dams suffered a schedule overrun.
9. Actual implementation schedule was on average 44% (or 2.3 years) higher than the estimate with a median of 27% (or 1.7 years) as shown in Fig. 5. Like cost overruns, the evidence is overwhelming that implementation schedules are systematically biased towards underestimation (Mann–Whitney–Wilcoxon $U=29{,}161$, $p<0.01$); the magnitude of schedule underestimation (i.e. schedule slippage) is larger than the error of schedule overestimation ($p<0.01$).
10. Graphing the dams' schedule overruns also reveals a fat tail as shown in Fig. 5, albeit not as fat as the tail of cost overruns. Costs are at a higher risk of spiraling out of control than schedules.
11. There is less variation in schedule overruns across regions than cost overruns. Large dams built everywhere take significantly longer than planners forecast. North America with a 27% mean schedule overrun is the best performer. A non-parametric comparison using a Wilcoxon test ($p=0.01$) suggests that projects in South Asia have significantly greater schedule overruns ($M=83\%$) than rest of the world taken as a whole ($M=42\%$). We investigate this further with a multilevel model below (Model 3).
12. There is no evidence for schedule estimates to have improved over time.

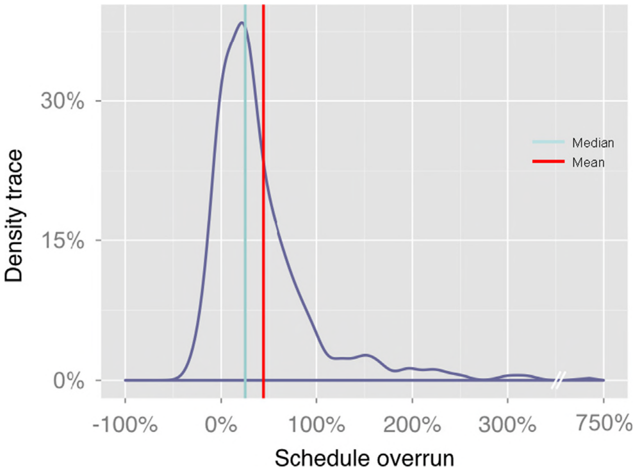


Fig. 5. Density trace of schedule slippage ($N=239$) with the median and mean.

We tested whether implementation schedules and project scale are related. A preliminary univariate analysis, which makes no attempts to take into account any covariates, shows that increase in the scale of a dam, e.g., measured as estimated cost of construction, increases the absolute actual implementation schedule required exponentially ($R^2=0.13$, $F=36.4$, $p<0.01$). Large scale is intimately linked with the long-term (see Model 2 below). The actual implementation schedule, reported here, does not take into the account lengthy lead times in preparing the projects. Dams require extensive technical and economic feasibility analysis, social and environmental impact studies, and political negotiations. The actual implementation cycles are far longer than the average of about 8.6 years, as shown in our data, that it takes to build a dam. These lengthy implementation schedules suggest that the benefits of large dams (even assuming that large dam generate benefits as forecasted) do not come “online” quickly enough. The temporal mismatch between when users need specific benefits and when these benefits come online is not to be downplayed (Ansar et al., 2012). Alternative investments that can bridge needs quickly, without tremendous time lags, are preferable to investments with a long lead-time and hence duration risk (Luehrman, 1998; Copeland and Tufano, 2004).

3.3. Multilevel regression analysis of cost and schedule performance

Means, standard deviations, and correlations of the variables used in the multilevel regressions are shown in Table 2.

We fitted multilevel regression models with projects nested by country as a second level to incorporate within-country correlation. The models were fitted using the “lme” procedure in the “nlme” package in R software. This function fits a linear mixed-effects model in the formulation described in Laird and Ware (1982) but allowing for nested random effects. The within-group errors are allowed to be correlated and/or have unequal variances. We found it necessary to transform variables to remove excessive skewness as noted in Table 2. Using stepwise variable selection, we are not only able to fit explanatory models for cost and overruns and estimated duration but also practicably parsimonious models for predicting them.

Table 3 summarizes the results from multilevel model examining predictors of cost overruns (Model 1). Model 1 identifies

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

Table 2
Descriptive statistics and correlations ($N=245$).

Variable	Mean	S.D.	1	2	3	4	5	6	7	8	9
1. Cost Overrun[a]	2.0	3.6									
2. Schedule slippage[a]	1.5	0.7	0.17**								
3. Estimated schedule (months)[b]	73.1	33.8	−0.16*	0.23**							
4. Actual schedule (months)[b]	102.7	55.7	−0.27**	−0.43**	0.76**						
5. Year—decision to build	1971.1	13.2	−0.02	0.05	−0.21**	−0.25**					
6. Year—completion	1979.6	12.7	−0.14*	−0.10	0.03	0.08	0.94**				
7. Project type dummy	0.8	0.4	−0.14*	0.08	0.10	0.02	−0.02	−0.02			
8. Democracy dummy	0.4	0.5	0.00	−0.14*	0.16*	0.20**	−0.45**	−0.38**	0.00		
9. Estimated cost (USD MM 2010 constant)[b]	699.6	1215.5	−0.03	0.09	0.48**	0.37**	0.02	0.13*	0.37**	−0.04	
10. Actual cost (USD MM 2010 constant)[b]	1462.2	4032.5	−0.38**	0.02	0.50**	0.43**	0.02	0.17**	0.38**	−0.03	0.93**
11. Height of dam wall (m)[c]	77.3	51.6	−0.10	0.10	0.26**	0.17**	0.10	0.16*	0.34**	−0.03	0.51**
12. Installed hydropower capacity (MW)[b]	487.0	1255.3	−0.16*	0.19**	0.22**	0.08	0.13*	0.16*	0.69**	−0.14*	0.59**
13. Length of dam wall (m)[b]	1364.1	2061.9	−0.12	−0.07	0.25**	0.30**	−0.19**	−0.08	−0.07	0.08	0.37**
14. Tunnel length (m)[b]	3500.0	7869.5	0.13	−0.12	−0.04	0.16	−0.06	−0.01	−0.23	0.05	0.11
15. Manufactures unit value index CAGR[d]	6.0	5.4	−0.01	−0.03	−0.25**	−0.18**	−0.12	−0.18**	0.08	−0.08	−0.13
16. GDP (nominal USD B)[b]	1221.1	253.4	−0.05	0.25**	0.36**	0.17*	0.29**	0.37**	−0.13	0.13	0.19*
17. Per capita income (2000 constant USD)[b]	4132.8	5198.6	0.23**	0.15*	0.11	0.01	−0.37**	−0.40**	−0.07	0.48**	−0.07
18. Long-term inflation (%)[b]	17%	0.2	−0.29**	0.04	−0.09	−0.11	0.22**	0.19**	0.24**	−0.37**	0.13*
19. Forex depreciation (%)[e]	18%	70.3	−0.30**	−0.04	0.03	0.00	0.29**	0.29**	0.16*	−0.20**	0.21**
20. South Asia dummy	0.1	0.3	−0.25**	−0.18**	0.17**	0.26**	−0.04	0.07	−0.06	0.20**	0.11
21. North America dummy	0.2	0.4	0.28**	0.06	0.21**	0.13*	−0.57**	−0.55**	−0.09	0.52**	0.06
Variable	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
11. Height of dam wall (m)	0.51**										
12. Installed hydropower capacity (MW)	0.60**	0.47**									
13. Length of dam wall (m)	0.38**	0.03	0.13								
14. Tunnel length (m)	−0.01	0.05	−0.22	−0.18							
15. Manufactures Unit Value Index CAGR	−0.12	−0.08	−0.02	0.02	−0.02						
16. GDP (nominal USD)	0.19*	0.10	0.09	0.04	−0.29	−0.31**					
17. Per capita income (2000 constant USD)	−0.14*	−0.08	−0.11	−0.02	−0.09	−0.01	0.29**				
18. Long-term inflation (%)	0.22**	0.06	0.33**	0.07	−0.41*	0.15*	−0.03	−0.24**			
19. Forex	0.29**	0.09	0.29**	−0.02	−0.37*	−0.16*	0.00	−0.26**	0.64**		
20. South Asia dummy	0.19**	0.08	−0.03	0.20**	NA	−0.09	−0.01	−0.46**	−0.10	0.11	
21. North America dummy	−0.03	−0.10	−0.16*	0.19**	NA	−0.18**	0.33**	0.60**	−0.44**	−0.31**	−0.15*

[a] One over ($1/x$) transformed.
[b] Log transformed.
[c] Sq. rt. ($\sqrt{x}$).
[d] Cb rt. ($\sqrt[3]{x}$).
[e] $x^{0.25}$ transformed to remove excess skewness for regression analysis and to calculate correlations.
** $p<0.01$.
* $p<0.05$.

Table 3
Model 1—Significant variables for cost accuracy for large dam projects (constant local currency).

Variable	Regression coefficient	Standard error	*t*-Stat	2-Tailed significance
Intercept	1.402	0.185	7.560	0.000
Log estimated duration (months)	−0.100	0.041	−2.424	0.016
Log of country's long-term inflation rate (%)	−0.085	0.029	−2.930	0.005

Note: Dependent variable is cost forecast accuracy, which is the estimated/actual cost ratio (i.e. $1/x$ of the cost overrun to remove excessive skewness), based on 239 observations. Since the dependent variable in Model 1 is the inverse of the cost overrun a negative sign on the coefficients of both significant variables suggests that an increase in the estimated duration or long-term inflation rate increases the cost overrun.

the estimated implementation schedule and the long-term inflation rate in the country in which the project is built as highly significant variables. An increase in estimated duration of one year contributes to an increase in cost overrun of approx. 5–6 percentage points depending on the country whilst holding the inflation rate constant (see Fig. A1). Note that an R-squared measure, which is customary to report for single-level regressions as explained proportion of variance, cannot be applied to multilevel models (Recchia, 2010).[5] The usual diagnostics, based upon the model residuals, were satisfactory.

The first finding in Model 1 is that the larger the estimated implementation schedule the higher the cost overrun ($p=0.016$), with all other things being equal, is particularly noteworthy for two reasons.

First, Model 1 suggests that planners' forecasting skills decay the longer in the future they are asked to project the risks facing a large dam. Material information about risks, for example, related to geology, prices of imports, exchange rates, wages, interest rates, sovereign debt, environment, only reveal in future shaping episode to which decision-makers are "blind" ex ante (Flyvbjerg and Budzier, 2011). We discuss some qualitative case examples to illustrate this statistical result and its broader implications in the next section.

Second, preliminary analysis had suggested that estimated implementation schedules depend on the scale of a planned

---

[5] Recchia (2010, p. 2) explains further why a *R*-squared measure cannot be used for a multilevel model. A single-level model "includes an underlying assumption of residuals that are independent and identically distributed. Such an assumption could easily be inappropriate in the two[or multi]-level case since there is likely to be dependence among the individuals that belong to a given group. For instance, it would be difficult to imagine that the academic achievements of students in the same class were not somehow related to one another". Also see Kreft and Leeuw (1998) and Goldstein (2010).

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

investment—i.e. bigger projects take longer to build. Support of this preliminary result was found by fitting a multilevel model (Model 2) that examines the predictors of estimated implementation schedule. Model 2 shows that height ($p=0.02$), installed capacity (MW) ($p=0.02$), and length ($p=0.04$) of the dam wall are significant variables associated with the estimated implementation schedule. The effect of these covariates can be seen from the coefficients in Table 4: a greater height, installed capacity, or length contribute to longer implementation schedules. We interpret Model 2 as follows. Estimated implementation schedule acts not only as a temporal variable but also as a surrogate for scalar variables such as wall height (which is also highly correlated with installed capacity). The larger the dam, the longer the estimated implementation schedule, and the higher the cost overrun.

Taken together, the multilevel models for cost overruns and estimated schedule suggest that longer time horizons and increasing scale are underlying causes of risk in investments in large hydropower dam projects.

The second finding in Model 1 is that the higher the long-term inflation rate of the host country the higher the cost overrun suffered by a dam ($p=0.02$). The long-term inflation rate was calculated by fitting a linear model to the log of the time series of the GDP deflator index of each country. The slope of this fitted line can be interpreted as the annual average growth rate of the log inflation for each country. This slope is a different constant for each country with some countries such as Brazil with a considerably higher long-term inflation rate, and hence greater propensity to cost overruns, than China or the United States. Moreover, this slope is stable in the short-run (it takes years of high or low inflation to change this slope) and hence our estimate can be assumed to be a reliable predictor. Recall that the cost overrun is being measured in constant terms (i.e. with the effects of inflation removed); yet Model 1 suggests that the inflation trajectory of a country, which we interpret as a surrogate of the overall macroeconomic management, is an important risk when making durable investments. The multilevel model finally suggests that once country specific factors have been taken into account the factor that drives cost overrun is the planning horizon.

Finally, we fit a multilevel model (Model 3) to examine predictors of schedule overruns. Model 3 identifies the following significant variables: whether or not a country is a democracy; the per capita income of the country in 2000 constant USD in the year of the decision to build; the planned installed capacity (MW); and planned length of the dam wall (meters). Avid dam building countries in South Asia, at various stages of democratic maturity, have also one of the poorest schedule performances in building dams. We controlled for this fact by including a dummy variable for South Asia in the model as a covariate with an interaction effect with the democracy dummy. Democracy in South Asia is significant in explaining schedule overruns. The South Asia dummy, however, does not come out to be significant. The effect of these covariates and the interaction effect can be seen in Table 5.

First, democracies' forecasts about implementation schedules of large dams are systematically more optimistic than autocracies even after controlling for systematically higher schedule overruns in India and Pakistan. The size of the coefficient is large suggesting that political process has profound impact on the schedule slippage. We tested whether democracies take longer than autocracies to build large dams by fitting a model to explain the actual implementation schedule (Model 4). Model 4, summarized in Table 6, shows that effects of political regime on the actual schedule are not significant. In other words, while democracies do not take longer to build large dams than autocracies in absolute terms, democracies appear to be more optimistic. Given its vast scope, we defer a further investigation of this important result to a future inquiry. We note, however, that theories of delusion and deception in the planning of large infrastructure projects (Flyvbjerg et al., 2009) would interpret this as evidence of ex ante political intent among democratically elected politicians to present a rosier picture about large dams than they know the case to be.

Second, countries with a higher per capita income in constant 2000 USD in the year of decision to build tend to have lower

Table 4
Model 2—Significant variables for estimated construction schedule for large dam projects (months).

Variable	Regression coefficient	Standard error	*t*-Stat	2-Tailed significance
Intercept	3.444	0.197	17.464	0.000
Sq rt of dam wall height (m)	0.029	0.012	2.414	0.017
Log of dam wall length (m)	0.058	0.027	2.153	0.033
Log of hydropower installed capacity (MW)	0.016	0.007	2.141	0.034

Note: Dependent variable is log of the estimated construction schedule, based on 239 observations.

Table 6
Model 4—Significant variables for estimated construction schedule for large dam projects (months).

Variable	Regression coefficient	Standard error	*t*-Stat	2-Tailed significance
Intercept	−17.712	6.401	−2.767	0.007
Log of dam wall length (m)	0.105	0.029	3.567	0.001
Year of actual project completion	0.011	0.003	3.358	0.001

Note: Dependent variable is log of the actual construction schedule, based on 239 observations.

Table 5
Model 3—Significant variables for schedule slippage for large dam projects.

Variable	Regression coefficient	Standard error	*t*-Stat	2-Tailed significance
Intercept	0.405	0.163	2.483	0.014
Democracy dummy[a]	−0.134	0.055	−2.439	0.016
Log of country's per capita income in year of decision to build (constant USD)	0.065	0.019	3.334	0.001
Log of dam wall length (m)	−0.027	0.013	−2.081	0.039
Log of hydropower installed capacity (MW)	0.018	0.006	3.207	0.002
South Asia dummy	0.211	0.113	1.874	0.066
Democracy in South Asia interaction effect	−0.239	0.113	−2.114	0.036

Note: Dependent variable is $1/x$ of the actual/estimated schedule ratio, based on 239 observations.

[a] Dummy based on the Polity2 variably of Polity IV regime index. Score of +10 to +6=democracy; score of +5 to −10=autocracy.

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

schedule overruns than countries with lower per capita income. We concur with the interpretation of Bacon and Besant-Jones (1998, p. 325) that "the best available proxy for most countries is [the] country-per-capita income…[for] the general level of economic support that a country can provide for the construction of complex facilities". This result suggests that developing countries in particular, despite seemingly the most in need of complex facilities such as large dams, ought to stay away from bites bigger than they can chew.

Third, the evidence appears to be contradictory with respect to scale. While a greater dam wall length contributes to a higher schedule overrun, a higher MW installed capacity has the opposite effect. Model 3 in Table 5 shows that the size of coefficients for the two significant variables related to physical scale—i.e. Log of dam wall length (m) and Log of hydropower installed capacity (MW)—is approximately the same but with the opposite sign.[6]

In attempting to interpret this result our conjecture is as follows. Dam walls are bespoke constructions tied to the geological and other site-specific characteristics. In contrast, installed capacity is manufactured off-site in a modular fashion. For example, the 690 MW installed capacity of the recently completed Kárahnjúkar project in Iceland was delivered with six generating units of identical design (6 × 115 MW). We propose that project components that require onsite construction, e.g. dam wall, are more prone to schedule errors than components manufactured off-site, e.g. generation turbines. Project designs that seek to reduce the bespoke and onsite components in favor of greater modular and manufactured components may reduce schedule uncertainty.

This conjecture is supported by Model 4 in Table 6, which shows that the actual construction schedule, in absolute terms, is significantly increased with an increase in the length of dam wall. In contrast, MW installed capacity does not have an effect on the absolute actual construction schedule suggesting that construction schedules are more sensitive to on-site construction than to components manufactured in factories. Note that lower installed capacity does not necessarily equate with a smaller dam. For example, it is not rare for a large multipurpose dam to have a low MW installed capacity when, for instance, the dam is primarily being used for irrigation or flood management purposes.

4. Qualitative case examples and policy propositions

The statistical results reported in the preceding sections show that cost and schedule estimates of large dams are severely and systematically biased below their actual values. While it is beyond the scope of this paper to discuss wider theoretical implications, the evidence presented here is consistent with previous findings that point to twin problems that cause adverse outcomes in the planning and construction of large and complex facilities such as large hydropower dams: (1) biases inherent in human judgment (delusion) and (2) misaligned principal-agent relationships or political incentives (deception) that underlie systematic forecasting errors. In the context of large dams, we argue that large scale and longer planning time horizons exacerbate the impact of these twin problems. We now present a few qualitative examples of risks large dams typically face to illustrate the statistical results reported above. We jointly draw on the statistical analyses and qualitative analyses to distill propositions of immediate relevance to policy.

Globally, experts' optimism about several risk factors contribute to cost overruns in large dams. For example, the planning documents for the Itumbiara hydroelectric project in Brazil recognized that the site chosen for the project was geologically unfavorable. The plan optimistically declared, "the cost estimates provide ample physical contingencies [20% of base cost] to provide for the removal of larger amounts [of compressible, weak, rock] if further investigations show the need" (World Bank, 1973). This weak geology ended up costing +96% of the base cost in real terms. Itumbiara's case is illustrative of a broader problem. Even though geological risks are anticipatable there is little planners can do to hedge against it. For example, exhaustive geological investigation for a large dam can cost as much as a third of the total cost (Hoek and Palmieri, 1998); at which point still remains a considerable chance of encountering unfavorable conditions that go undetected during the *ex ante* tests (Goel et al., 2012).

Policy proposition 1. *Energy alternatives that rely on fewer site-specific characteristics such as unfavorable geology are preferable.*

Similarly, in the Chivor hydroelectric project in Colombia, the planning document was upbeat that there will be no changes in the exchange rate between the Colombian Peso and the U.S. dollar during the construction period (1970–1977) stating, "No allowance has been made for possible future fluctuations of the exchange rate. This approach is justified by recent experience in Colombia where the Government has been pursuing the enlightened policy of adjusting [policy] quickly to changing conditions in the economy" (World Bank, 1970). In fact, the Colombian currency depreciated nearly 90% against the U.S. dollar as shown in Fig. 6.

Since over half the project's costs covers imported inputs, this currency depreciation caused a 32% cost overrun in real Colombian Peso terms. Currency exposure arises when the inputs required to build a project are denominated in one currency but the outputs in another, or vice versa. The outputs of dams, such as electricity, are denominated in the local currency. Similarly, any increases in tax receipts a dam may enable for the host government also accrue in local currency. A large portion of inputs to build a dam, particularly in developing countries, however, constitute imports paid for in USD. Since the USD liabilities also have to eventually be

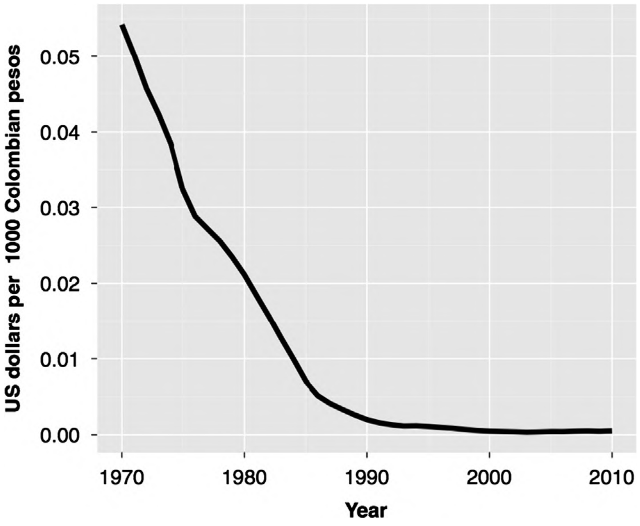


Fig. 6. Depreciation of the Colombian Peso 1970–2010.

---

[6] Note that the dependent variable in Model 3 is forcast accuracy, the inverse of schedule overrun (i.e. $1/x$ of the schedule overrun or Estimated/Actual schedule). Thus a negative sign on the Log of dam wall length (m) suggests that an increase in wall length decreases the inverse of the schedule overrun. In other words, increase in wall length increases schedule overrun.

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

paid in local currency, currency exposure consistently proves to be a fiscal hemorrhage for large projects.

Policy proposition 2. *Energy alternatives that rely on fewer imports or match the currency of liabilities with the currency of future revenue are preferable.*

Although, following convention, our cost analysis excludes the effects of inflation, planners ought not to ignore the risks of "unanticipated inflation" (Pickrell, 1992, p. 164). Episodes of hyperinflation in Argentina, Brazil, Turkey, and Yugoslavia caused staggering nominal cost overruns, e.g. 7-times initial budget for Brazil's Estreito dam (1965–1974), or 110-times initial budget for Yugoslavia's Visegrad dam (1985–1990), which had to be financed with additional debt. Effects of unanticipated inflation magnify the longer it takes to complete a project. For example, during the planning phase of Pakistan's Tarbela dam, it was assumed that inflation would not have a signification impact on the project's costs. The appraisal report wrote: "A general contingency of 7½% has been added in accordance with normal practice for works of this size and duration" (World Bank, 1968). The project, launched in 1968, was meant to start full commercial operation in 1976, but the opening was delayed until 1984. Actual cumulative inflation in Pakistan during 1968–1984 was 380%; the actual cost of the dam in nominal terms nearly four times the initial budget. In the case of Tarbela, unanticipated inflation was "a product of delays in a project's construction timetable and a higher-than expected inflation rate" (Pickrell, 1992, p. 164). For our reference class, 8 out of 10 large dams came in late with an average delay of 2.3 years. Moreover, forecasters expected the annual inflation rate to be 2.5% but it turned out to be 18.9% (averages for the entire sample). Large dams have a high propensity to face unanticipated inflation.

Policy proposition 3. *The best insurance against creeping inflation is to reduce the implementation schedule to as short a horizon as possible. Energy alternatives that can be built sooner and with lower risk of schedule overruns, e.g. through modular design, are preferable.*

Large dams are typically financed from public borrowing. While our calculations exclude debt-servicing, cost overruns increase the stock of debt but also the recurring financing costs that can further escalate if interest rates go up. The optimistic risk assessments of the costs of large dams are consistent with "explosive growth of Third World debt" (Bulow and Rogoff, 1990; Mold, 2012). For example, the actual cost of Tarbela dam, the majority of which was borrowed from external sources, amounted to 23% of the increase in Pakistan's external public debt stock between 1968 and 1984; or 12% for Colombia's Chivor dam (1970–1977) as shown in Table 7.

These case examples reinforce the essential message of our statistical results: bigger projects entail uncontrollable risks, which even when anticipatable cannot be adequately hedged. We do not directly negate the presence of economies of scale or learning curves—i.e. declining average cost per unit as output increases. Instead our argument is that any economies of scale embedded in large scale are being acquired for a disproportionately increased exposure to risk that can cause financial impairment. Companies and countries with insufficient capacity to absorb adverse outcomes of big bets gone awry often face financial ruin.

Table 7
Total stock of public net external debt (USD current, MM).

Year	Colombia	Pakistan
1968		3252.4
1970	1296.6	
1977	2699.6	
1984		9692.8
Debt increase over the implementation schedule	1403.0	6440.5
Cost of mega-dam over the relevant period (USD current MM)	Chivor dam	Tarbela dam
	168.7	1497.90
Cost of dam as percentage of debt increase	12.0%	23.2%

Policy proposition 4. *Energy alternatives that do not constitute a large proportion of the balance sheet of a country or a company are preferable. Similarly, policymakers, particularly in countries at lower levels of economic development, ought to avoid highly leveraged investments denominated in a mix of currencies.*

5. Forecasting the actual costs and schedules using reference class forecasting (RCF)

As discussed in the methods section, the third step of the "outside view" or RCF techniques is to compare a specific venture with the reference class distribution, in order to establish the most likely outcome for the specific venture. Thus if systematic errors in the forecasts generated using the "inside view" of previous ventures are found, decision-makers should apply an uplift or downlift to the "inside view" forecast in order to generated a de-biased "outside view" forecast. For example, empirical literature has established that rail projects suffer a cost overrun of 45% on average (Flyvbjerg, 2008; also see Table 8). The 50th percentile cost overrun for rail projects is 40% and the 80th percentile is 57%. Based on these findings, RCF techniques suggest that decision-makers ought to apply a 57% uplift to the initial estimated budget in order to obtain 80% certainty that the final cost of the project would stay within budget (Flyvbjerg, 2008, p. 16). If decision-makers were more risk tolerant then they could apply a 40% uplift to the initial estimated budget but then there will remain a 50% chance that the proposed project might exceed its budget.

In line with the RCF techniques, the third and final step of our investigation on dams was to derive a good predictor of cost and schedule overruns for proposed large dams based on the distributional information of the reference class. This predictor serves to "correct" the systematically biased *ex ante* cost and schedule estimates by adjusting them upwards by the average cost or schedule overrun (see Kahneman and Tversky, 1979b; Flyvbjerg, 2006, 2008).

First, using traditional RCF (Flyvbjerg, 2006, 2008), we traced the empirical distribution of cost and schedule overruns of large dams. Second, we use multilevel Models 1 and 3, described above, for predicting cost and schedule overruns. Models 1 and 3 prove to be practicably parsimonious models for two reasons: First both models are fitted with variables known *ex ante*. Second, both models were successfully fitted with only a few significant variables making it practicable to collect the data needed to make a prediction. For example, Model 1 on cost overruns has only two significant variables—estimate schedule and the long-term inflation rate of the host country. Data on both these variables is readily available for any proposed large dam making it possible to predict the cost overrun before construction begins. We illustrate the usefulness of our predictive models with an example below.

With respect to cost overruns, using traditional RCF (Flyvbjerg, 2006, 2008), we find that if planners are willing to accept a 20% risk of a cost overrun, the uplift required for large dams is +99% (i.e. ~double experts' estimates) as seen in Fig. 7; and +176%

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

Table 8
Comparing large dams with other infrastructure asset classes.

Category	**Types of projects**	**Mean cost overrun**	Applicable capital expenditure optimism bias uplifts (constant prices)	
			50th percentile	**80th percentile**
Roads	Motorway, trunk roads, local road, bicycle facilities, pedestrian facilities, park and ride, bus lane schemes, guided buses	20%	15%	32%
Rail	Metro, light rail, guided buses on tracks, conventional rail, high speed rail	45%	40%	57%
Fixed links	Bridges, tunnels	34%	23%	55%
Building projects	Stations, terminal buildings		4–51%[a]	
Standard civil engineering			3–44%[a]	
Non-standard civil engineering			6–66%[a]	
Mining projects		14%[b]		
Thermal power plants		6%[c]		
Large dam projects	Large hydropower, large irrigation, flood control, multipurpose dams	**96%**	26%	**99%**
Nuclear power plants		207%[d]		109–281%[d]

[a] Based on Mott MacDonald (2002).
[b] Based on Bertisen and Davis (2008).
[c] Based on Bacon and Besant-Jones (1998, p.321), included for an approximate comparison purposes only, reference class probability distribution not available.
[d] Based on Schlissel and Biewald (2008, p.8) review of the U.S. Congressional Budget Office (CBO) data from Energy Information Administration, Technical Report DOE/EIA-0485 (January 1, 1986).

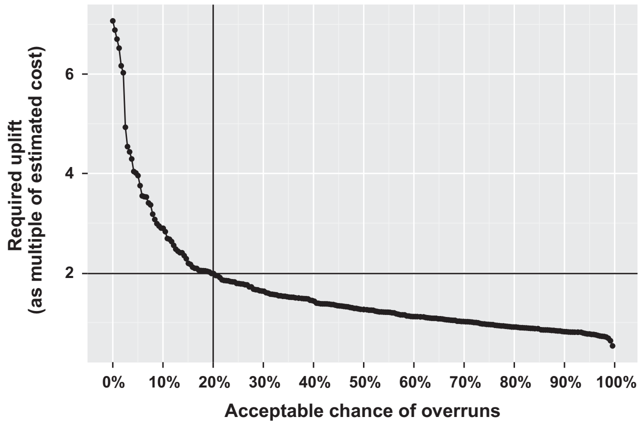


Fig. 7. Required uplift for large dam projects as function of the maximum acceptable level of risk for cost overrun, constant local currency terms ($N=245$).

including unanticipated inflation. If planners are willing to accept a 50–50 chance of a cost overrun, the uplift required is 26% (32% outside North America).

In terms of cost overruns, Fig. 7 also illustrates that large dams are one of the riskiest asset classes for which valid and reliable data are available. Compare, for example, Fig. 7 with reference class forecasts previously conducted for rail, road, tunnel, or bridge projects (Flyvbjerg, 2006, 2008) also summarized in Table 8.

Second, using our multilevel Model 1 we were able to derive predictions for cost overrun (in constant local currency) and schedule overrun respectively.

Experts estimate, for instance, that Pakistan's Diamer-Bhasha dam, whose construction began shortly after the 2010 floods, will cost PKR 894 billion (~USD12.7B in 2008 prices and exchange rates and about 9% of Pakistan's 2008 GDP) (WAPDA, 2011). The dam is forecasted to take 10 years from 2011 and become operational in 2021. Using our first approach, the reference class forecast for cost overruns suggests that planners need to budget PKR 1,788B (USD25.4B) in real terms to obtain 80% certainty of not exceeding the revised budget. Including the effects of unanticipated inflation the required budget is PKR 2,467B (USD35.0B) or about 25% of Pakistan's 2008 GDP. A future sovereign default in Pakistan owing to this one mega-dam is not a remote possibility.

Using our second approach, our multilevel Model 1 predicts that given the 10 year estimated duration and a long-term inflation rate of about 8% the expected (average) cost overrun of a large dam in Pakistan will be 44% (PKR 1,288B or USD 18.3B). Combining the two methods, a conservative estimate for the cost overrun on the Diamer-Bhasha dam is 44% at which point there remains a 4 in 10 chance of the revised budget being exceeded. Note, however, that if a dam of dimensions similar to Diamer-Bhasha were being built in the US, Model 1 predicts that it would only suffer a cost overrun of 16%, which the much larger US economy could absorb without any lasting damage.

We applied a similar two-pronged forecast of schedule slippage. Using our first approach, the reference class forecast for schedule slippage suggests that planners for large dams around the world need to allow for a 66% schedule overrun to achieve 80% certainty that the project will be completed within the revised implementation schedule. Since Diamer-Bhasha is expected to take 10 years to build (2011–2021), planners need to adjust their schedule estimate upwards to nearly 17 years (i.e. an actual opening date of 2028). Using our second approach, our multilevel Model 3 predicts that given that the dam's final decision to build was made in Pakistan by a democratically elected government, when the per capita income was USD 497 in 2000 constant dollars, a dam wall length of 998 m, and an installed capacity of 4500 MW, the expected outcome is a 60% schedule overrun. Thus, using either approach, Diamer-Bhasha can be expected to only open in 2027 when there remains a 20% risk of further delay. Pakistan is facing an energy crisis today (Kessides, 2011). A dam that brings electricity is 2027 will be a little late in coming.

Note, however, that if a dam of dimensions similar to Diamer-Bhasha were being built in the US (with its high per capita income of approximately USD 38,000), Model 3 predicts that it would face a schedule slippage of a mere 0.05%. Recall that per capita income

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

is a useful proxy for the economic support that a country can provide for the construction of complex facilities. This suggests that rich and not developing countries best attempt very large energy projects, such as large dams. Even so, richer countries should also consider alternatives and should adopt the risk management measures of the outside view illustrated here to choose prudently among energy alternatives.

Using their "inside" cost estimates, the net present benefits to cost ratio of the dam according to experts is 1.43 (WAPDA, 2011). Even assuming experts' calculations about potential benefits are accurate, although this is a doubtful assumption, the de-biased cost forecasts require an uplift of 44–99% in constant prices suggest that the benefits to cost ratio will be below one. The Diamer-Bhasha dam is a non-starter in Pakistan. This is without even discussing potential effects of inflation and interest rates, potential social and environmental costs, and opportunity cost Pakistan could earn by committing such vast amount of capital to more prudent investments.

Our reference class forecasting techniques suggests that other proposed large dam projects such as Belo Monte, Myitsone, or the Gilgel Gibe III among many others in early planning stages are likely to face large cost and schedule overruns seriously undermining their economic viability. Large dams also exert an opportunity cost by consuming scarce resources that could be deployed to better uses, sinking vast amounts of land that could have yielded cash flows and jobs from agricultural, timber, or mineral resources. Risks related to dam safety, environment, and society further undermine viability of large dams. Decision-makers are advised to carefully stress test their proposed projects using the risk management techniques of the outside view proposed here before committing resources to them.

The outside view techniques applied to large dams above have broader application in energy policy by helping public agencies (e.g. national planning and finance ministries, power and water authorities), private entrepreneurs and investors a framework to improve upfront selection among alternatives. The problems of cost and schedule overrun are not unique to large hydropower dams. Preliminary research suggests that other large-scale power projects using nuclear, thermal, or wind production technologies face similar issues. Our research of large hydropower projects reveals that there is a serious dearth of valid and reliable data on the risk profiles of actually completed energy projects across the board. Much of the data in existing literature are drawn from surveys and interviews of dubious validity. At times, interest groups, seeking to promote a particular kind of scale or technology, also report distorted data. There is thus an urgent need to empirically document, in a comprehensive global database, the risk profiles of energy infrastructure assets of large, medium, and small scales across production technologies. For example, comparing the likely actual cost, schedule, and production volumes of a large hydropower dam project versus an on-site combined heat and power generator.

We propose that prior to making any energy investment, policy makers consult a valid and reliable "outside view" or "reference class forecast" (RCF) that can predict the outcome of a planned investment of a particular scale or production technology based on actual outcomes in a reference class of similar, previously completed, cases. Rigorously applying reference class forecasting to energy investments at various scales and production technologies will yield the following contributions:

- Create transparency on risk profiles of various energy alternatives, from not only the perspective of financial cost and benefit but also environmental and social impact—hard evidence is a counter-point to experts' and promoters' oft-biased inside view.

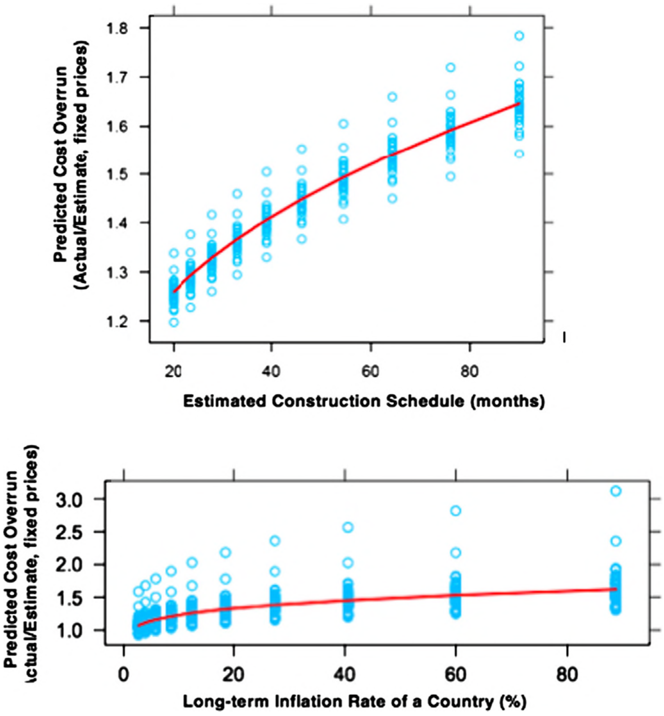


Fig. A1. .

- Improve resource allocation through outside-in view to estimate costs, benefits, time, and broader impacts such as greenhouse gas emissions incurred in building a project and emission created or averted once a project becomes operational.

A comprehensive global dataset that can create such transparency on risk profiles of energy alternatives does not yet exist. We have sought to bridge this precise gap by providing impartial evidence on large hydropower dam projects. As a venue for further research we hope valid and reliable data on the actual cost, schedules, benefits, and impacts of other production technologies will become available to enable comparative analysis with novel implications for theory and practice.

Acknowledgments

This work was supported by funding from Oxford University's Centre of Major Programme Management at Saïd Business School. The authors are grateful to the United Nations Environment Programme (UNEP) and Messrs. Jeremy Bird and Bruce Aylward who made access to data from the World Commission of Dams possible. We also thank Mr. John Besant-Jones at the World Bank for advice on data collection. Professor Lord Robert May of Oxford, Professor Nassim Taleb, Professor Jim Hall, and two anonymous referees provided helpful comments on earlier versions of the paper.

Appendix A. Visual representation of Model 1 (reported in Table 3)

See Fig. A1.

References

Ansar, A., Flyvbjerg, B., Budzier, A., 2013. Do "Hard States" Make Better Decisions? Incentives versus Biases. Unpublished manuscript. Saïd Business School, University of Oxford, 42 pp.

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

Ansar, A., Flyvbjerg, B., Budzier, A., 2012. Big is Fragile: The Effects of Temporal Uncertainty and Sunk Costs on the Outcomes of Big Ventures. Unpublished manuscript. Saïd Business School, University of Oxford, 58 pp.

Bacon, R.W., Besant-Jones, J.E., 1998. Estimating construction costs and schedules: experience with power generation projects in developing countries. Energy Policy 26, 317–333.

Bakker, K., 1999. The politics of hydropower: developing the Mekong. Political Geogr. 18, 209–232.

Bertisen, Jasper, Davis, Graham A., 2008. Bias and Error in Mine Project Capital Cost Estimation. The Eng. Economist 53 (2), 118–139.

Billington, D.P., Jackson, D.C., 2006. Big Dams of the New Deal Era: A Confluence of Engineering and Politics. University of Oklahoma Press.

Buehler, R., Griffin, D., Ross, M., 1994. Exploring the“ planning fallacy”: Why people underestimate their task completion times. J. Pers. Soc. Psychol. 67, 366.

Bulow, J., Rogoff, K., 1990. Cleaning up third world debt without getting taken to the cleaners. J. Econ. Perspect. 4, 31–42.

Copeland, T., Tufano, P., 2004. A real-world way to manage real options. Harv. Bus. Rev. 82, 90–99.

Dawes, R.M., Faust, D., Meehl, P.E., 1989. Clinical versus actuarial judgment. Science 243, 1668–1674.

Duflo, E., Pande, R., 2007. Dams. Q. J. Econ. 122, 601–646.

Federal Transit Administration (FTA), 2003. Predicted and Actual Impacts of New Starts Projects: Capital Cost, Operating Cost and Ridership Data. U.S. Department of Transportation, Washington, DC.

Flyvbjerg, B., 2003. Delusions of success: comment on Dan Lovallo and Daniel Kahneman. Harv. Bus. Rev. 81, 121–122.

Flyvbjerg, B., 2006. From Nobel Prize to project management: getting risks right. Project Manage. J. 37, 5.

Flyvbjerg, B., 2008. Curbing optimism bias and strategic misrepresentation in planning: reference class forecasting in practice. Eur. Plann. Stud. 16, 3–21.

Flyvbjerg, B., 2009. Survival of the unfittest: why the worst infrastructure gets built—and what we can do about it. Oxford Rev. Econ. Policy 25, 344.

Flyvbjerg, B., Bruzelius, N., Rothengatter, W., 2003. Megaprojects and Risk: An Anatomy of Ambition. Cambridge Univ Press, Cambridge.

Flyvbjerg, B., Budzier, A., 2011. Why your IT project may be riskier than you think. Harv. Bus. Rev. 89, 23–25.

Flyvbjerg, B., COWI, 2004. Procedures for dealing with optimism bias in transport planning: Guidance document 2006–09 [WWW Document]. URL goo.gl/lLiNj.

Flyvbjerg, B., Holm, M., Buhl, S., 2005. How (in) accurate are demand forecasts in public works projects. J. Am. Plann. Assoc. 71, 131–146.

Flyvbjerg, B., Holm, M.S., Buhl, S., 2002. Underestimating costs in public works projects: error or lie? J. Am. Plann. Assoc. 68, 279–295.

Flyvbjerg, B., Garbuio, M., Lovallo, D., 2009. Delusion and deception in large infrastructure projects. Calif. Manage. Rev., 51.

Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, Cambridge.

Gilovich, T., Griffin, D., Kahneman, D., 2002. Heuristics and Biases: The Psychology of Intuitive Judgment. Cambridge University Press.

Goel, R.K., Singh, B., Zhao, J., 2012. Underground Infrastructures: Planning, Design, and Construction. Butterworth-Heinemann, Waltham, MA.

Goldstein, H., 2010. Multilevel Statistical Models, 4th ed. Wiley.

Hoek, E., Palmieri, A., 1998. Geotechnical risks on large civil engineering projects. In: Proceedings of the 8th Congres IAEG, pp. 79–88.

Hufschmidt, M.M., Gerin, J., 1970. Systematic errors in cost estimates for public investment projects. In: Margolis, J. (Ed.), The Analysis of Public Output. Columbia University Press, New York.

ICOLD, 2010. International. Commission on Large Dams [WWW Document]. URL ⟨http://www.icold-cigb.org/⟩.

IEA, 2011. Policies and Scenarios [WWW Document]. URL ⟨http://www.iea.org/publications/scenariosandprojections/⟩.

Kahneman, D., 1994. New challenges to the rationality assumption. J. Inst. Theor. Econ. 150, 18–36.

Kahneman, D., 2011. Thinking, Fast and Slow, 1st ed. Farrar, Straus and Giroux, New York.

Kahneman, D., Lovallo, D., 1993. Timid choices and bold forecasts: a cognitive perspective on risk taking. Manage. Sci., 17–31.

Kahneman, D., Tversky, A., 1979a. Prospect theory: an analysis of decision under risk. Econometrica, 263–291.

Kahneman, D., Tversky, A., 1979b. Intuitive prediction: biases and corrective procedures. In: Makridakis, S., Wheelwright, S.C. (Eds.), Studies in the Management Sciences: Forecasting. North Holland, Amsterdam.

Kessides, I.N., 2011. Chaos in Power: Pakistan's Electricity Crisis.

Kreft, I.G.G., Leeuw, J.d.e., 1998. Introducing Multilevel Modeling. Sage Publications Ltd.

Laird, N.M., Ware, J.H., 1982. Random-effects models for longitudinal data. Biometrics, 963–974.

Lovallo, D., Kahneman, D., 2003. Delusions of success. Harv. Bus. Rev. 81, 56–63.

Luehrman, T.A., 1998. Strategy as a portfolio of real options. Harv. Bus. Rev. 76, 89–101.

McCully, P., 2001. Silenced Rivers: the Ecology and Politics of Large Dams. Zed Books, London.

Merewitz, L., 1973. How do Urban Rapid Transit Projects Compare in Cost Estimating Experience. Institute of Urban & Regional Development, University of California.

Merrow, E.W., McDonnell, L., Argüden, R.Y., Corporation, R., 1988. Understanding the Outcomes of Megaprojects: A Quantitative Analysis of Very Large Civilian Projects. Rand Corp.

Mold, A., 2012. Will it all end in tears? Infrastructure spending and African development in historical perspective. J. Int. Dev. 24, 237–254.

Mott MacDonald, 2002. Review of Large Public Procurement in the UK, Study for HM Treasury. HM Treasury, London. www.hm-treasury.gov.uk/d/7(3).pdf.

Nilsson, C., Reidy, C.A., Dynesius, M., Revenga, C., 2005. Fragmentation and flow regulation of the world's large river systems. Science 308, 405–408.

Pickrell, D.H., 1989. Urban Rail Transit Projects: Forecast Versus Actual Ridership and Costs (No. PB90148693), Office of Grants Management, Urban Mass Transportation Administration. Transportation System Center, U.S. Department of Transportation, Cambridge, MA.

Pickrell, D.H., 1992. A desire named streetcar: fantasy and fact in rail transit planning. American Planning Association. J. Am. Plann. Assoc. 58, 158–176.

Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., Woodhouse, G., Draper, D., Langford, I., Lewis, T., 2009. A User's Guide to MLwiN. University of London, Institute of Education, Centre for Multilevel Modelling.

Recchia, A., 2010. R-squared measures for two-level hierarchical linear models using SAS. J. Stat. Softw., 32.

Richter, B.D.a., Postel, S.b., Revenga, C.c., Scudder, T.d., Lehner, B.e., Churchill, A.f., Chow, M., 2010. Lost in development's shadow: the downstream human consequences of dams. Water Alternatives 3, 14–42.

Salazar, J.G., 2000. Damming the child of the ocean: the Three Gorges project. J. Environ. Dev. 9, 160–174.

Scudder, T., 2005. The Future of Large Dams: Dealing with Social, Environmental, Institutional and Political Costs. Earthscan, London.

Schlissel, D., Biewald, B., 2008. Nuclear Power Plant Construction Costs. Synapse Energy Economics, Inc.

Singh, S., 2002. Taming the Waters: The Political Economy of Large Dams in India. Oxford University Press, New Delhi.

Sovacool, B.K., Bulan, L.C., 2011. Behind an ambitious megaproject in Asia: the history and implications of the Bakun hydroelectric dam in Borneo. Energy Policy 39, 4842–4859.

Sovacool, B.K., Cooper, C., 2013. The Governance of Energy Megaprojects: Politics, Hubris and Energy Security. Edward Elgar, Cheltenham, UK.

Stone, R., 2011. Hydropower. The legacy of the Three Gorges Dam. Science 333, 817.

Taleb, N.N., 2010. The Black Swan: The Impact of the Highly Improbable with a new section: “On Robustness and Fragility”, 2nd ed. Random House Trade Paperbacks.

Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: heuristics and biases. Science 185, 1124–1131.

Wachs, M., 1989. When planners lie with numbers. J. Am. Planning Assoc. 55, 476–479.

WAPDA, 2011. Diamer-Basha Dam Project Status as on 20 June 2011.

World Bank, 1968. Report and Recommendation of the President to the Executive Directors on a Proposed Loan for the Tarbela Project and on Proposed Extension of Closing Dates of Loan 266-PAK (Indus Basin Project), Pakistan (No. P-616). The World Bank, Washington, DC.

World Bank, 1970. Staff Appraisal Report (pu-31a) of the Chivor Hydroelectric Project, Interconexion Electrico S.A., Colombia (No. PU-31a). The World Bank, Washington, DC.

World Bank, 1973. Staff Appraisal Report (150a-BR) of the Itumbiara Hydroelectric Power Project, FURNAS, Brazil (No. 150a-BR). The World Bank, Washington, DC.

World Bank, 1996. Estimating Construction Costs and Schedules: Experience with Power Generation Projects in Developing Countries (No. WTP325). The World Bank, Washington, DC.

World Commission on Dams (WCD), 2000. Cross-Check Survey: Final Report. Cape Town, South Africa.

Ziv, G., Baran, E., Nam, S., Rodríguez-Iturbe, I., Levin, S.A., 2012. Trading-off fish biodiversity, food security, and hydropower in the Mekong River Basin. Proc. Natl Acad. Sci. 109, 5609.

Please cite this article as: Ansar, A., et al., Should we build more large dams? The actual costs of hydropower megaproject development. Energy Policy (2014), http://dx.doi.org/10.1016/j.enpol.2013.10.069

TAB 3

RISK.1721

Variability in Accuracy Ranges: A Case Study in the Canadian Hydropower Industry

John K. Hollmann, PE CCP CEP DRMP,
Raminder S. Bali, P.Eng, John M. Boots, P.Eng,
Chantale Germain, P.Eng,
Michel Guevremont, P.Eng PSP,
and Kelman Kai-man Ng, P.Eng

Abstract—This paper presents a case study of the variability in accuracy ranges for cost estimates in the Canadian hydropower industry. The study sought to improve the participants' understanding of risks and estimate accuracy for their hydropower projects of similar scope. The study team also sought to verify the theoretical accuracy curves identified in AACE International's Recommended Practice (RP) 69R-12: "Cost Estimate Classification System – As Applied in Engineering, Procurement, and Construction for the Hydropower Industry". The study team collected and analyzed actual and phased estimate cost data from 24 projects with actual costs from 50 million to 3.6 billion (2012$CAN) completed from 1974 to 2014. Greenfield, brownfield and revamp impoundment and hydropower generation facility projects from across Canada were included (power transmission projects were excluded.) The study found that the range bandwidth (uncertainty) in RP 69R-12 is understated. Further, because actual contingency estimates are biased too low, the actual range curves are biased very high relative to those in RP 69R-12. The accuracy ranges and the underestimation of contingency are similar for hydropower and process industry projects.

Table of Contents

List of Figures

List of Tables

Introduction

Accuracy is a measure of how a cost estimate differs from the final actual outcome. Risk analysis provides forecasts of how the final actual outcome may differ from the estimate (such as a base estimate or an amount approved for expenditure). Historical analysis helps us to understand the variability of accuracy and to improve our risk analysis practice [1]. This study is such an historical analysis.

Empirical estimate accuracy data has been researched for over 50 years [2]. In particular, the accuracy of process industry project estimates (e.g., oil and gas, chemical, mining, etc.) has been well documented [3]. Other studies have highlighted industry bias and misperceptions of the reality of estimate accuracy [4]. However, there has been a relative void in accuracy studies for hydropower projects with the notable exception of studies of World Bank funded projects; mostly in developing countries [5,6]. This study of the accuracy of estimates for the well developed Canadian hydropower industry will help fill a gap in our understanding of the hydropower industry.

In addition, this study was needed to help verify the applicability of the theoretical accuracy depiction presented in Figure 1 of AACE International's new Recommended Practice 69R-12: "Cost Estimate Classification System – As Applied in Engineering, Procurement, and Construction for the Hydropower Industry" [7]. The questions in regard to that RP were "does Figure 1 in RP 69R-12 reflect real accuracy ranges?" and if not, "how can we assure that this depiction does not feed bias in stakeholder expectations?" (Figure 1 from RP 69R-12 is reproduced below):

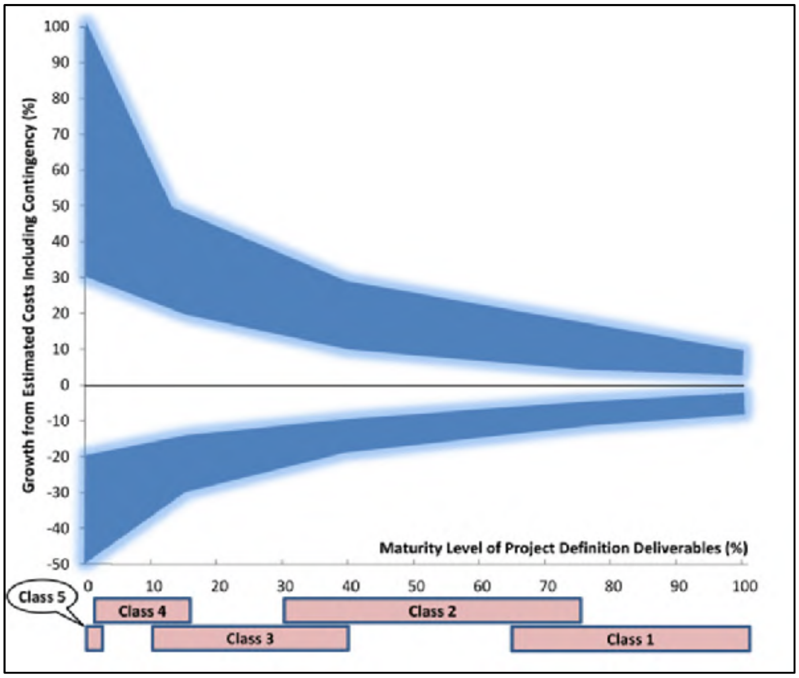


Figure 1 – Example of the Variability in Accuracy Ranges for a Hydropower Industry Estimate (Figure 1 from AACE International RP 69R-12; copied with permission)

Background on the Study

RP 69R-12 referenced above resulted from a multi-year effort led by a team of Canadian hydropower cost experts. The initial RP goal was to document the defining scope deliverables and their expected status to support hydropower project estimates of each Class. After publication of the RP, the next step was to verify theoretical Figure 1 in the RP. To do this, the Canadian study team performed an empirical analysis. The project scope and cost details of the analysis are confidential to the team, but it is hoped that the reference information presented here will be useful for AACE International to improve RP 69R-12.

The Canadian study team collected estimated and actual project capital cost data from 24 projects with actual costs from $50 million to $3.6 billion (in 2012 $CAN) completed from 1974 to 2012. For each project, estimate data from each scope development phase was captured, resulting in data on 50 estimates. All projects had a Class 3 estimate, but some did not have Class 4 and/or 5. The project scopes included greenfield, brownfield and major revamp impoundment and power generation facilities on rivers across Canada. It excluded power transmission projects. Most of the projects were located in semi-remote areas and included camps, mass excavation, concrete and/or earth-filled impoundments and diversions, intake structures, penstocks, and power houses with turbine generation equipment. To minimize bias, the dataset represented all the recent major project data available to the participants regardless of whether the project cost outcome met company objectives.

Analysis Approach

The primary analytical methods used were descriptive statistics and multi-variable linear regression. The accuracy metric described by the statistics and the dependent variable of regression was the ratio of “base estimate/actual costs”. “Base estimates” exclude contingency, escalation and management reserves. This was used because the team wanted to understand how actual costs differed from the base so that they could improve future predictions of this difference (i.e., predict contingency required). The study also examined schedule duration estimate accuracy which is not included in this paper.

The estimate/actual cost ratio was used because it tends to be close to a normal distribution and hence is amenable to linear regression analysis. As will be discussed later, the more commonly considered actual/estimate inverse tends to be biased to the high side which makes regression analysis problematic.

The primary independent variables examined included (drivers of accuracy outcomes):

- Scope definition (i.e., Class)
- Cost content (e.g., % equipment or impoundment)
- Location/Company
- Proximity to populated areas

- Cost/Schedule Strategy (i.e., cost or schedule driven)
- Terrain/Site Conditions/Weather
- New Technology or Scale
- System Complexity
- Execution Complexity
- Primary Project Type (e.g., greenfield, revamp, etc.)
- Primary Construction Contract Type
- Owner PM System Maturity

To collect the data, the team developed a form that captured the following:

- General project characteristics
- High level "base" cost estimate breakdowns at each AACE Class (per 69R-12) plus contingency and escalation cost estimates for each
- Actual final cost
- Key planned and actual schedule milestones
- Scope change and risk event information

The actual cost data was normalized to the year of the respective estimate using the mid-point of spending approach (actual project cash flows were not available) [8]. The normalization price index used was derived from Statistics Canada indices for the sell price of non-residential construction projects. Also, cost changes due to business scope change were adjusted out (costs resulting from a change to a basic premise of the estimate such as generation capacity or throughput.) None of the projects were observed to have experienced a catastrophic risk event.

The primary variable (risk driver) of interest was the level of scope definition. Not all projects had data for estimates of each AACE Class as can be seen in the following number of valid observations:

- Class 3: 21 (a group of 4 projects in a program were combined into one)
- Class 4: 17
- Class 5: 12

Data for 3 projects was excluded because extreme age and/or duration raised questions as to the validity of the normalization. This sample size was considered adequate to gain useful insight as to the relationship of accuracy and Class, but not enough to gain deep understanding of the impact on accuracy of any but the most dominant of the other independent variables. The linear regression was performed using Microsoft Excel® with an add-on package called Analyse-it® that provides additional modeling, diagnostic and graphical capabilities.

Findings for Accuracy Range by Class: Descriptive Statistics

Table 1 shows the dataset statistics for accuracy. Figure 2 depicts the same data fitted to lognorm distributions. The probability values ("p-value" is the level of confidence expressed as a percentage of values that will be less than that shown) in the table are calculated using the Excel "Norminv" function applied to the base estimate/actual data, and then converted to the traditional actual/base estimate ratio format (i.e., >1 means the actual cost was more than the base estimate.) This method of inferring the population distribution from a sample is consistent with the method described in AACE International RP 42R-08 and supported by process industry research that indicates that estimate/actual data (as opposed to its inverse of actual/estimate) is more or less normally distributed [1].

As an example of how to interpret this, if the ratio for Class 3 at p50 is 1.24, that indicates that 24% contingency would be needed to achieve a 50 percent confidence of underrunning. Note the high side skewing (e.g. the Class 5 p90 of 3.01 is much further from the mean than the p10 value). Recall that these values exclude escalation and business scope change.

Actual/Base Estimate	Class 3	Class 4	Class 5
number of observations	*21*	*17*	*12*
Mean	1.24	1.40	1.79
p90	1.63	2.09	3.01
p50	1.24	1.40	1.79
p10	0.99	1.06	1.27

Table 1 – Dataset Cost Estimate Accuracy Metrics (Actual/Base Estimate)

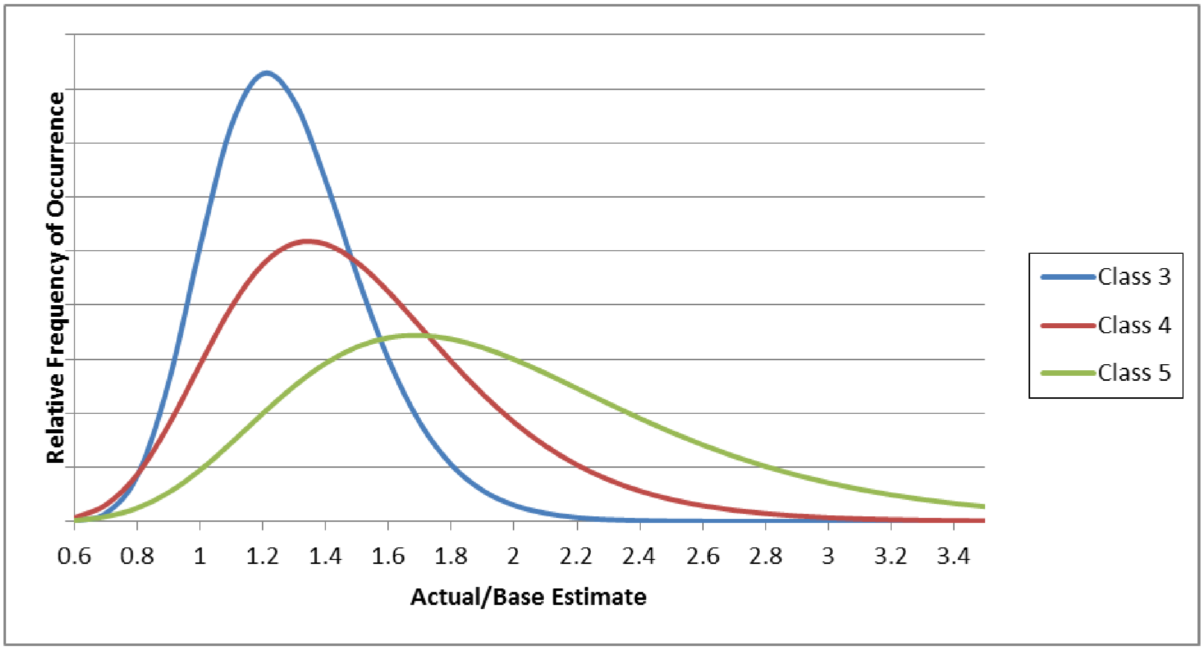


Figure 2 – Dataset Actual/Base Estimate Metrics Fitted to Lognorm Distributions

Comparison of Findings to Other Studies and AACE RP69R-12

Statistically speaking, considering sample sizes and data quality, this study's accuracy ranges are comparable to those reported for the process and infrastructure industries [3 and 4] as well as hydropower projects funded by the World Bank [5]. Table 2 summarizes the results of these studies. It was assumed that funding estimates in studies [4] and [5] were based on about Class 4 scope definition because general industry front-end planning is assumed to be less defined than planning at the companies in this study and at the clientele of Independent Project Analysis, Inc. (IPA). Note that this study's values were adjusted downward from Table 1 to reflect the accuracy relative to the estimate *including* contingency (i.e., the funded amount) which is the data shown in most published studies. The contingencies added to this study's Class 3, 4 and 5 base estimates were 10%, 12% and 15% respectively which correspond to typical contingencies applied at the time.

This Study, Canadian Hydro	**Class 3**	**Class 4**	**Class 5**
p90	53%	97%	186%
p50	14%	28%	64%
p10	-11%	-6%	12%
IPA Inc., Process Industry [3]; p10/p90 approximated from histogram illustration			
p90	40%	70%	200%
p50	1%	5%	38%
p10	-15%	-15%	-15%
Hollmann, Process Industry [4] average of meta-analysis			
p90		70%	
p50		21%	
p10		-9%	
Merrow, Hydro [5] Mean & Std Dev Reported; Normal distribution assumed below)			
p90 (assuming normal)		65%	
Mean		24%	
p10 (assuming normal)		-17%	

Table 2 – Comparison of Accuracy Studies (% Overrun of Estimate Including Contingency)

When comparing results in respect to RP 69R-12, one must consider two points of comparison. The first is the bandwidth or span of the range (i.e., p90 minus p10.) The other is the absolute value of a high or low range. Figure 3 shows this study's results superimposed on the RP 69R-12 Figure 1. This study's range spans are somewhat wider (more uncertain) than the worst case spans in the RP. For example, the worst case span for Class 5 in the RP is 150% (100 – <50>) while the span for Class 5 in this study is 174% (186 – 12.) The high and low absolute range values indicate strong contingency under-estimation bias. Note that all the projects in the study were greater than $50 million (in 2012 $CAN); the findings may not apply to small projects where estimating practices often differ. [4]

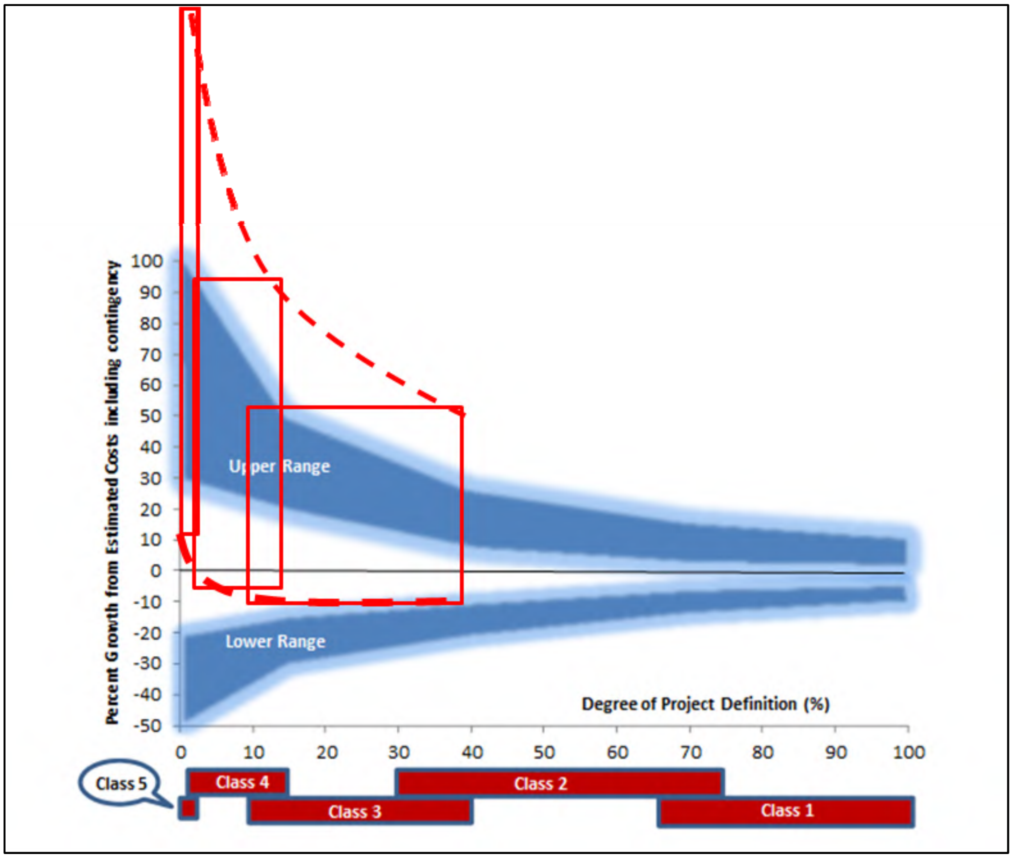


Figure 3 – RP 69R-12 Figure 1 with p10/p90 Study Data Superimposed

Comparison of Contingency Estimates to Actual Cost Growth

The projects in this study allowed only 10-15% contingency on average, even for Class 5 estimates. These contingencies appear to reflect a strong industry optimism bias. For example, a 2012 white paper by the United States Society on Dams, suggested "An overall contingency of *as much as* 50 percent is appropriate" on the conceptual engineer's estimate (i.e., Class 5); however, this *maximum* ("as much as") contingency allowance is much less than this study's *mean* cost growth of 79% at Class 5 [9].

Contingencies (or combinations of contingency and management reserve) of 24, 40 and 79% at p50, excluding business scope change and escalation, are suggested by this study for Class 3, 4 and 5 estimates respectively for projects of average risks. If these contingencies had been included in the study projects, their actual range outcome would look similar to but wider than the worst case of RP 69R-12 Figure 1. The authors are not recommending that these or any other contingency values be assigned arbitrarily; contingency should always be based on risk analyses. However, if a company's risk analyses regularly result in 10-15% contingency and narrow ranges, it is likely that risks and their impacts are not being identified or quantified properly and/or optimism bias is controlling.

Findings for Other Risk Drivers from Regression Analysis

An attempt was made to quantify the impacts of systemic risks other than the level of scope definition. To do this, the data from only the Class 3 estimates was examined. Class 3 is usually the basis for full funding decisions and hence of utmost importance to the stakeholders. Each independent variable (risk driver) was tested alone and in various combinations using Excel with Analyze-it.

A regression model quantifying the cost growth for Class 3 estimates was developed that had an R2 of 0.66. Because the dataset had only 21 observations, the actual model is not shown here to avoid any misuse (findings may not be generally applicable,) but narratively speaking, the following variables appear to be significant systemic risk drivers:

- Proximity: The greater the distance of the project from a large population center, the greater the cost growth. Given the effect of distance on material and labor availability and conditions, this seems rational.
- Size: Larger projects had *less* cost growth. This may be due to larger projects being the sum of parts with highs and lows that balance out, and/or smaller project estimates may be small because of bias towards lowering base costs resulting in greater cost growth.
- Months Execution Duration: Longer projects had more cost growth despite normalizing for escalation. This may reflect the fact that risks often drive both cost and schedule increases rather than a causal correlation. However, the more time that passes, the more chance that there will systemic changes in the social, political, regulatory and other environments.
- % Equipment in Estimate: The greater the proportion of equipment, the less the cost growth. Excluding scope change, most estimators would agree that major equipment is less subject to risk and uncertainty than labor and bulk material costs, particularly for impoundments subject to geologic risks.

The analysis above was repeated with Estimate Class added as an independent variable. This could serve the participant's as a rudimentary parametric model for systemic risk analysis [1]. While the model is confidential, it can be said that the model coefficients for Class (the level of scope definition) are consistent with the relative range values for each Class in Table 1 and Figure 1. The level of scope definition is clearly the predominant systemic risk driver.

Conclusion

This study of the variability in accuracy ranges for cost estimates in the Canadian hydropower industry suggests that the actual cost uncertainty is a bit greater than the worst case theoretical depiction of accuracy in Figure 1 of RP 69R-12. The study indicated that risks are much greater than being estimated; contingencies of 24, 40 and 79 percent were indicated for Class 3, 4 and 5 estimates respectively on average. Our study shows that the contingency and reserves estimated were lower than what were required. However, the Canadian hydro industry experience is similar to that of other process industry projects, as well as of hydropower projects in other regions funded by the World Bank.

Using the data from the study, the participants developed a simple parametric risk analysis tool for systemic risks in which the level of scope definition as the dominant risk driver. This emphasizes the importance of doing disciplined Class 3 scope definition prior to full funds authorization if cost predictability is a goal. The Canadian hydro study team will recommend that AACE's Cost Estimating Technical Committee consider improvements to Figure 1 and related content in RP 69R-12 to reflect the findings of this study. The conclusions are applicable to other process related industries, and therefore this paper may encourage improvements in other estimate classification recommended practices.

References

1. AACE International, "*Risk Analysis and Contingency Determination Using Parametric Estimating*", Recommended Practice 42R-08, AACE International, Morgantown, WV, (latest revision).
2. Linder, Stefan, "*Fifty years of Research on Accuracy of Capital Expenditure Project Estimates: A Review of the Findings and their Validity*", Center for Research in Controllership and Management, WHU - Otto Beisheim Graduate School of Management, Vallendar, Germany, April 2005.
3. Ogilvie, Alexander, Robert A. Brown Jr., Fredrick P. Biery, and Paul Barshop, *"Quantifying Estimate Accuracy and Precision for the Process Industries: A Review of Industry Data",* Cost Engineering, November/December 2012, AACE International, Morgantown, WV.
4. Hollmann, John K., *"Estimate Accuracy-Dealing with Reality",* Cost Engineering, November/December 2012, AACE International, Morgantown, WV.
5. Merrow, Edward W. and Brett R Schroeder, "*Understanding the Costs and Schedule of Hydroelectric Projects*", 1991 AACE International Transactions, AACE International, Morgantown, WV, 1991.
6. Bacon, Robert W, John E. Besant-Jones, and Jamshid Heidarian, "*Estimating Construction Costs and Schedules; Experience with Power Generation Projects in Developing Countries"*, World Bank, Washington DC, 1996.
7. AACE International, "*Cost Estimate Classification System – As Applied in Engineering, Procurement, and Construction for the Hydropower Industry*", Recommended Practice 69R-12, AACE International, Morgantown, WV, (latest revision).
8. AACE International, "*Escalation Principles and Methods Using Indices*", Recommended Practice 58R-10, AACE International, Morgantown, WV, (latest revision).
9. United States Society on Dams, *"Guidelines for Construction Cost Estimating for Dam Engineers and Owners" (white paper),* US Society on Dams, May 2012.

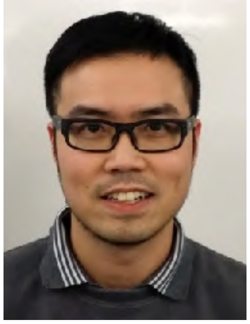
John K. Hollmann, PE CCP CEP DRMP
Validation Estimating, LLC
jhollmann@validest.com

Raminder S. Bali, P.Eng
Manitoba Hydro
rbali@hydro.mb.ca

John M. Boots, P.Eng
BC Hydro
john.boots@bchydro.com

Chantale Germain, P.Eng
Hydro Quebec
germain.chantale@hydro.qc.ca

Michel Guevremont, P.Eng PSP
Hydro Quebec
guevremont.michel@hydro.qc.ca



Kelman Kai-man Ng, P.Eng
BC Hydro
kelman.ng@bchydro.com

TAB 4

AACE International Recommended Practice No. 18R-97

COST ESTIMATE CLASSIFICATION SYSTEM – AS APPLIED IN ENGINEERING, PROCUREMENT, AND CONSTRUCTION FOR THE PROCESS INDUSTRIES

TCM Framework: 7.3 – Cost Estimating and Budgeting

Rev. November 29, 2011

Note: As AACE International Recommended Practices evolve over time, please refer to www.aacei.org for the latest revisions.

Contributors:

Peter Christensen, CCE (Author)
Larry R. Dysert, CCC CEP (Author)
Jennifer Bates, CCE
Jeffery J. Borowicz, CCE CEP PSP
Peter R. Bredehoeft, Jr. CEP
Robert B. Brown, PE
Dorothy J. Burton
Robert C. Creese, PE CCE
John K. Hollmann, PE CCE CEP
Kenneth K. Humphreys, PE CCE
Donald F. McDonald, Jr. PE CCE PSP
C. Arthur Miller
Todd W. Pickett, CCC CEP
Bernard A. Pietlock, CCC CEP
Wesley R. Querns, CCE
Don L. Short, II CEP
H. Lance Stephenson, CCC
James D. Whiteside, II PE

AACE® International Recommended Practice No. 18R-97

COST ESTIMATE CLASSIFICATION SYSTEM – AS APPLIED IN ENGINEERING, PROCUREMENT, AND CONSTRUCTION FOR THE PROCESS INDUSTRIES

TCM Framework: 7.3 – Cost Estimating and Budgeting

November 29, 2011

PURPOSE

As a recommended practice of AACE International, the *Cost Estimate Classification System* provides guidelines for applying the general principles of estimate classification to project cost estimates (i.e., cost estimates that are used to evaluate, approve, and/or fund projects). The *Cost Estimate Classification System* maps the phases and stages of project cost estimating together with a generic project scope definition maturity and quality matrix, which can be applied across a wide variety of process industries.

This addendum to the generic recommended practice (17R-97) provides guidelines for applying the principles of estimate classification specifically to project estimates for engineering, procurement, and construction (EPC) work for the process industries. This addendum supplements the generic recommended practice by providing:

- a section that further defines classification concepts as they apply to the process industries; and
- a chart that maps the extent and maturity of estimate input information (project definition deliverables) against the class of estimate.

As with the generic recommended practice, an intent of this addendum is to improve communications among all of the stakeholders involved with preparing, evaluating, and using project cost estimates specifically for the process industries.

The overall purpose of this recommended practice is to provide the process industry definition deliverable maturity matrix which is not provided in 17R-97. It also provides an approximate representation of the relationship of specific design input data and design deliverable maturity to the estimate accuracy and methodology used to produce the cost estimate. The estimate accuracy range is driven by many other variables and risks, so the maturity and quality of the scope definition available at the time of the estimate is not the sole determinate of accuracy; risk analysis is required for that purpose.

This document is intended to provide a guideline, not a standard. It is understood that each enterprise may have its own project and estimating processes and terminology, and may classify estimates in particular ways. This guideline provides a generic and generally acceptable classification system for process industries that can be used as a basis to compare against. This addendum should allow each user to better assess, define, and communicate their own processes and standards in the light of generally-accepted cost engineering practice.

INTRODUCTION

For the purposes of this addendum, the term process industries is assumed to include firms involved with the manufacturing and production of chemicals, petrochemicals, and hydrocarbon processing. The common thread among these industries (for the purpose of estimate classification) is their reliance on process flow diagrams (PFDs) and piping and instrument diagrams (P&IDs) as primary scope defining documents. These documents are key deliverables in determining the degree of project definition, and thus the extent and maturity of estimate input information.

Estimates for process facilities center on mechanical and chemical process equipment, and they have significant amounts of piping, instrumentation, and process controls involved. As such, this addendum may apply to portions of other industries, such as pharmaceutical, utility, metallurgical, converting, and similar industries. Specific addendums addressing these industries may be developed over time.

This addendum specifically does not address cost estimate classification in non-process industries such as commercial building construction, environmental remediation, transportation infrastructure, hydropower, “dry” processes such as assembly and manufacturing, “soft asset” production such as software development, and similar industries. It also does not specifically address estimates for the exploration, production, or transportation of mining or hydrocarbon materials, although it may apply to some of the intermediate processing steps in these systems.

The cost estimates covered by this addendum are for engineering, procurement, and construction (EPC) work only. It does not cover estimates for the products manufactured by the process facilities, or for research and development work in support of the process industries. This guideline does not cover the significant building construction that may be a part of process plants.

This guideline reflects generally-accepted cost engineering practices. This addendum was based upon the practices of a wide range of companies in the process industries from around the world, as well as published references and standards. Company and public standards were solicited and reviewed, and the practices were found to have significant commonalities. These classifications are also supported by empirical process industry research of systemic risks and their correlation with cost growth and schedule slip[8].

COST ESTIMATE CLASSIFICATION MATRIX FOR THE PROCESS INDUSTRIES

	Primary Characteristic	***Secondary Characteristic***		
ESTIMATE CLASS	**MATURITY LEVEL OF PROJECT DEFINITION DELIVERABLES** Expressed as % of complete definition	**END USAGE** Typical purpose of estimate	**METHODOLOGY** Typical estimating method	**EXPECTED ACCURACY RANGE** Typical variation in low and high ranges [a]
Class 5	0% to 2%	Concept screening	Capacity factored, parametric models, judgment, or analogy	L: -20% to -50% H: +30% to +100%
Class 4	1% to 15%	Study or feasibility	Equipment factored or parametric models	L: -15% to -30% H: +20% to +50%
Class 3	10% to 40%	Budget authorization or control	Semi-detailed unit costs with assembly level line items	L: -10% to -20% H: +10% to +30%
Class 2	30% to 75%	Control or bid/tender	Detailed unit cost with forced detailed take-off	L: -5% to -15% H: +5% to +20%
Class 1	65% to 100%	Check estimate or bid/tender	Detailed unit cost with detailed take-off	L: -3% to -10% H: +3% to +15%

Notes: [a] The state of process technology, availability of applicable reference cost data, and many other risks affect the range markedly. The +/- value represents typical percentage variation of actual costs from the cost estimate after application of contingency (typically at a 50% level of confidence) for given scope.

Table 1 – Cost Estimate Classification Matrix for Process Industries

Table 1 provides a summary of the characteristics of the five estimate classes. The maturity level of definition is the sole determining (i.e., primary) characteristic of Class. In Table 1, the maturity is roughly indicated by a % of complete definition; however, it is the maturity of the defining deliverables that is the determinant, not the percent. The specific deliverables, and their maturity, or status, are provided in Table 3. The other characteristics are secondary and are generally correlated with the maturity level of project definition deliverables, as discussed in the generic RP[1]. The characteristics are typical for the process industries but may vary from application to application.

This matrix and guideline outline an estimate classification system that is specific to the process industries. Refer to the generic estimate classification RP[1] for a general matrix that is non-industry specific, or to other addendums for guidelines that will provide more detailed information for application in other specific industries. These will provide additional information, particularly the project definition deliverable maturity matrix which determines the class in those particular industries.

Table 1 illustrates typical ranges of accuracy ranges that are associated with the process industries. Depending on the technical and project deliverables (and other variables) and risks associated with each estimate, the accuracy range for any particular estimate is expected to fall into the ranges identified (although extreme risks can lead to wider ranges).

In addition to the degree of project definition, estimate accuracy is also driven by other systemic risks such as:

- Level of non-familiar technology in the project.
- Complexity of the project.
- Quality of reference cost estimating data.
- Quality of assumptions used in preparing the estimate.
- Experience and skill level of the estimator.
- Estimating techniques employed.
- Time and level of effort budgeted to prepare the estimate.

Systemic risks such as these are often the primary driver of accuracy; however, project-specific risks (e.g. risk events) also drive the accuracy range[3].

Another way to look at the variability associated with estimate accuracy ranges is shown in Figure 1. Depending upon the technical complexity of the project, the availability of appropriate cost reference information, the degree of project definition, and the inclusion of appropriate contingency determination, a typical Class 5 estimate for a process industry project may have an accuracy range as broad as -50% to +100%, or as narrow as -20% to +30%.

Figure 1 also illustrates that the estimating accuracy ranges overlap the estimate classes. There are cases where a Class 5 estimate for a particular project may be as accurate as a Class 3 estimate for a different project. For example, similar accuracy ranges may occur for the Class 5 estimate of one project that is based on a repeat project with good cost history and data and the Class 3 estimate for another project involving new technology. It is for this reason that Table 1 provides ranges of accuracy range values. The accuracy range is determined through risk analysis of the specific project.

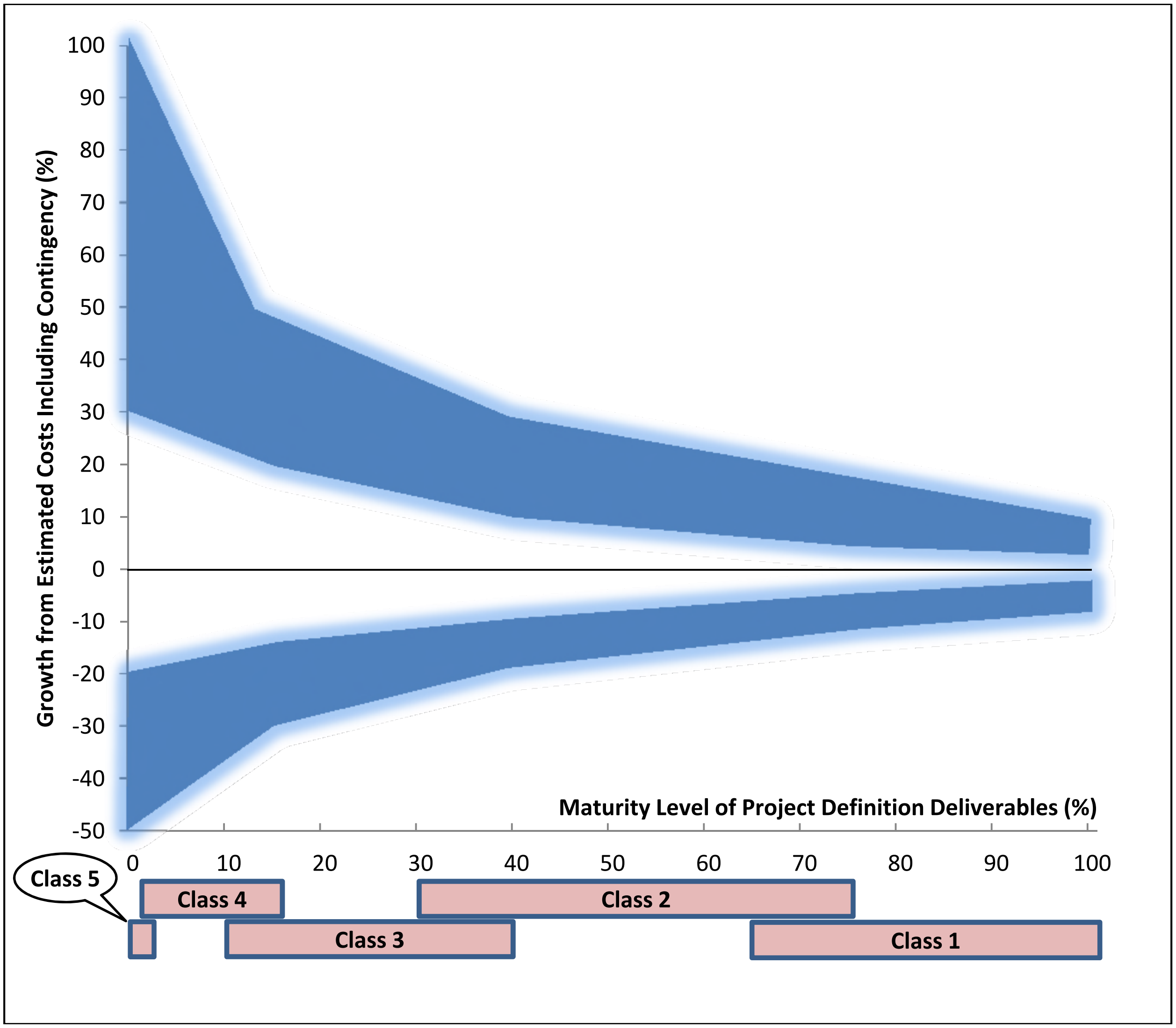


Figure 1 – Example of the Variability in Accuracy Ranges for a Process Industry Estimate

DETERMINATION OF THE COST ESTIMATE CLASS

The cost estimator makes the determination of the estimate class based upon the maturity level of project definition based on the status of specific key planning and design deliverables. The percent design completion may be correlated with the status, but the percentage should not be used as the Class determinate. While the determination of the status (and hence class) is somewhat subjective, having standards for the design input data, completeness and quality of the design deliverables will serve to make the determination more objective.

CHARACTERISTICS OF THE ESTIMATE CLASSES

The following tables (2a through 2e) provide detailed descriptions of the five estimate classifications as applied in the process industries. They are presented in the order of least-defined estimates to the most-defined estimates. These descriptions include brief discussions of each of the estimate characteristics that define an estimate class.

For each table, the following information is provided:

- **Description:** a short description of the class of estimate, including a brief listing of the expected estimate inputs based on the maturity level of project definition deliverables. The “minimum” inputs reflect the range of industry experience, but would not generally be recommended.
- **Maturity Level of Project Definition Deliverables (Primary Characteristic):** Describes a particularly key deliverable and a typical target status in stage-gate decision processes, plus an indication of approximate percent of full definition of project and technical deliverables. For the process industries, this correlates with the percent of engineering and design complete.
- **End Usage (Secondary Characteristic):** a short discussion of the possible end usage of this class of estimate.
- **Estimating Methodology (Secondary Characteristic):** a listing of the possible estimating methods that may be employed to develop an estimate of this class.
- **Expected Accuracy Range (Secondary Characteristic):** typical variation in low and high ranges after the application of contingency (determined at a 50% level of confidence). Typically, this represents about 80% confidence that the actual cost will fall within the bounds of the low and high ranges. The estimate confidence interval or accuracy range is driven by the reliability of the scope information available at the time of the estimate in addition to the other variables and risk identified above.
- **Alternate Estimate Names, Terms, Expressions, Synonyms:** this section provides other commonly used names that an estimate of this class might be known by. These alternate names are not endorsed by this Recommended Practice. The user is cautioned that an alternative name may not always be correlated with the class of estimate as identified in Tables 2a-2e.

CLASS 5 ESTIMATE	
Description: Class 5 estimates are generally prepared based on very limited information, and subsequently have wide accuracy ranges. As such, some companies and organizations have elected to determine that due to the inherent inaccuracies, such estimates cannot be classified in a conventional and systematic manner. Class 5 estimates, due to the requirements of end use, may be prepared within a very limited amount of time and with little effort expended—sometimes requiring less than an hour to prepare. Often, little more than proposed plant type, location, and capacity are known at the time of estimate preparation. **Maturity Level of Project Definition Deliverables:** Key deliverable and target status: Block flow diagram agreed by key stakeholders. 0% to 2% of full project definition. **End Usage:** Class 5 estimates are prepared for any number of strategic business planning purposes, such as but not limited to market studies, assessment of initial viability, evaluation of alternate schemes, project screening, project location studies, evaluation of resource needs and budgeting, long-range capital planning, etc.	**Estimating Methodology:** Class 5 estimates generally use stochastic estimating methods such as cost/capacity curves and factors, scale of operations factors, Lang factors, Hand factors, Chilton factors, Peters-Timmerhaus factors, Guthrie factors, and other parametric and modeling techniques. **Expected Accuracy Range:** Typical accuracy ranges for Class 5 estimates are -20% to -50% on the low side, and +30% to +100% on the high side, depending on the technological complexity of the project, appropriate reference information and other risks (after inclusion of an appropriate contingency determination). Ranges could exceed those shown if there are unusual risks. **Alternate Estimate Names, Terms, Expressions, Synonyms:** Ratio, ballpark, blue sky, seat-of-pants, ROM, idea study, prospect estimate, concession license estimate, guesstimate, rule-of-thumb.

Table 2a – Class 5 Estimate

CLASS 4 ESTIMATE	
Description: Class 4 estimates are generally prepared based on limited information and subsequently have fairly wide accuracy ranges. They are typically used for project screening, determination of feasibility, concept evaluation, and preliminary budget approval. Typically, engineering is from 1% to 15% complete, and would comprise at a minimum the following: plant capacity, block schematics, indicated layout, process flow diagrams (PFDs) for main process systems, and preliminary engineered process and utility equipment lists. **Maturity Level of Project Definition Deliverables:** Key deliverable and target status: Process flow diagrams (PFDs) issued for design. 1% to 15% of full project definition. **End Usage:** Class 4 estimates are prepared for a number of purposes, such as but not limited to, detailed strategic planning, business development, project screening at more developed stages, alternative scheme analysis, confirmation of economic and/or technical feasibility, and preliminary budget approval or approval to proceed to next stage.	**Estimating Methodology:** Class 4 estimates generally use stochastic estimating methods such as equipment factors, Lang factors, Hand factors, Chilton factors, Peters-Timmerhaus factors, Guthrie factors, the Miller method, gross unit costs/ratios, and other parametric and modeling techniques. **Expected Accuracy Range:** Typical accuracy ranges for Class 4 estimates are -15% to -30% on the low side, and +20% to +50% on the high side, depending on the technological complexity of the project, appropriate reference information, and other risks (after inclusion of an appropriate contingency determination). Ranges could exceed those shown if there are unusual risks. **Alternate Estimate Names, Terms, Expressions, Synonyms:** Screening, top-down, feasibility (pre-feasibility for metals processes), authorization, factored, pre-design, pre-study.

Table 2b – Class 4 Estimate

CLASS 3 ESTIMATE	
Description: Class 3 estimates are generally prepared to form the basis for budget authorization, appropriation, and/or funding. As such, they typically form the initial control estimate against which all actual costs and resources will be monitored. Typically, engineering is from 10% to 40% complete, and would comprise at a minimum the following: process flow diagrams, utility flow diagrams, preliminary piping and instrument diagrams, plot plan, developed layout drawings, and essentially complete engineered process and utility equipment lists. **Maturity Level of Project Definition Deliverables:** Key deliverable and target status: Piping and instrumentation diagrams (P&IDs) issued for design. 10% to 40% of full project definition. **End Usage:** Class 3 estimates are typically prepared to support full project funding requests, and become the first of the project phase control estimates against which all actual costs and resources will be monitored for variations to the budget. They are used as the project budget until replaced by more detailed estimates. In many owner organizations, a Class 3 estimate is often the last estimate required and could very well form the only basis for cost/schedule control.	**Estimating Methodology:** Class 3 estimates generally involve more deterministic estimating methods than stochastic methods. They usually involve predominant use of unit cost line items, although these may be at an assembly level of detail rather than individual components. Factoring and other stochastic methods may be used to estimate less-significant areas of the project. **Expected Accuracy Range:** Typical accuracy ranges for Class 3 estimates are -10% to -20% on the low side, and +10% to +30% on the high side, depending on the technological complexity of the project, appropriate reference information, and other risks (after inclusion of an appropriate contingency determination). Ranges could exceed those shown if there are unusual risks. **Alternate Estimate Names, Terms, Expressions, Synonyms:** Budget, scope, sanction, semi-detailed, authorization, preliminary control, concept study, feasibility (for metals processes) development, basic engineering phase estimate, target estimate.

Table 2c – Class 3 Estimate

CLASS 2 ESTIMATE	
Description: Class 2 estimates are generally prepared to form a detailed contractor control baseline (and update the owner control baseline) against which all project work is monitored in terms of cost and progress control. For contractors, this class of estimate is often used as the bid estimate to establish contract value. Typically, engineering is from 30% to 75% complete, and would comprise at a minimum the following: process flow diagrams, utility flow diagrams, piping and instrument diagrams, heat and material balances, final plot plan, final layout drawings, complete engineered process and utility equipment lists, single line diagrams for electrical, electrical equipment and motor schedules, vendor quotations, detailed project execution plans, resourcing and work force plans, etc. **Maturity Level of Project Definition Deliverables:** Key deliverable and target status: All specifications and datasheets complete including for instrumentation. 30% to 75% of full project definition. **End Usage:** Class 2 estimates are typically prepared as the detailed contractor control baseline (and update the owner control baseline) against which all actual costs and resources will now be monitored for variations to the budget, and form a part of the change management program.	**Estimating Methodology:** Class 2 estimates generally involve a high degree of deterministic estimating methods. Class 2 estimates are prepared in great detail, and often involve tens of thousands of unit cost line items. For those areas of the project still undefined, an assumed level of detail takeoff (forced detail) may be developed to use as line items in the estimate instead of relying on factoring methods. **Expected Accuracy Range:** Typical accuracy ranges for Class 2 estimates are -5% to -15% on the low side, and +5% to +20% on the high side, depending on the technological complexity of the project, appropriate reference information, and other risks (after inclusion of an appropriate contingency determination). Ranges could exceed those shown if there are unusual risks. **Alternate Estimate Names, Terms, Expressions, Synonyms:** Detailed control, forced detail, execution phase, master control, engineering, bid, tender, change order estimate.

Table 2d – Class 2 Estimate

CLASS 1 ESTIMATE	
Description: Class 1 estimates are generally prepared for discrete parts or sections of the total project rather than generating this level of detail for the entire project. The parts of the project estimated at this level of detail will typically be used by subcontractors for bids, or by owners for check estimates. The updated estimate is often referred to as the current control estimate and becomes the new baseline for cost/schedule control of the project. Class 1 estimates may be prepared for parts of the project to comprise a fair price estimate or bid check estimate to compare against a contractor's bid estimate, or to evaluate/dispute claims. Typically, overall engineering is from 65% to 100% complete (some parts or packages may be complete and others not), and would comprise virtually all engineering and design documentation of the project, and complete project execution and commissioning plans. **Maturity Level of Project Definition Deliverables:** Key deliverable and target status: All deliverables in the maturity matrix complete. 65% to 100% of full project definition. **End Usage:** Generally, owners and EPC contractors use Class 1 estimates to support their change management process. They may be used to evaluate bid checking, to support vendor/contractor negotiations, or for claim evaluations and dispute resolution. Construction contractors may prepare Class 1 estimates to support their bidding and to act as their final control baseline against which all actual costs and resources will now be monitored for variations to their bid. During construction, Class 1 estimates may be prepared to support change management.	**Estimating Methodology:** Class 1 estimates generally involve the highest degree of deterministic estimating methods, and require a great amount of effort. Class 1 estimates are prepared in great detail, and thus are usually performed on only the most important or critical areas of the project. All items in the estimate are usually unit cost line items based on actual design quantities. **Expected Accuracy Range:** Typical accuracy ranges for Class 1 estimates are -3% to -10% on the low side, and +3% to +15% on the high side, depending on the technological complexity of the project, appropriate reference information, and other risks (after inclusion of an appropriate contingency determination). Ranges could exceed those shown if there are unusual risks. **Alternate Estimate Names, Terms, Expressions, Synonyms:** Full detail, release, fall-out, tender, firm price, bottoms-up, final, detailed control, forced detail, execution phase, master control, fair price, definitive, change order estimate.

Table 2e – Class 1 Estimate

ESTIMATE INPUT CHECKLIST AND MATURITY MATRIX

Table 3 maps the extent and maturity of estimate input information (deliverables) against the five estimate classification levels. This is a checklist of basic deliverables found in common practice in the process industries. The maturity level is an approximation of the completion status of the deliverable. The completion is indicated by the following letters.

- **None (blank):** development of the deliverable has not begun.
- **Started (S):** work on the deliverable has begun. Development is typically limited to sketches, rough outlines, or similar levels of early completion.
- **Preliminary (P):** work on the deliverable is advanced. Interim, cross-functional reviews have usually been conducted. Development may be near completion except for final reviews and approvals.
- **Complete (C):** the deliverable has been reviewed and approved as appropriate.

November 29, 2011

	ESTIMATE CLASSIFICATION				
	CLASS 5	**CLASS 4**	**CLASS 3**	**CLASS 2**	**CLASS 1**
MATURITY LEVEL OF PROJECT DEFINITION DELIVERABLES	**0% to 2%**	**1% to 15%**	**10% to 40%**	**30% to 75%**	**65% to 100%**
General Project Data:					
Project Scope Description	General	Preliminary	Defined	Defined	Defined
Plant Production/Facility Capacity	Assumed	Preliminary	Defined	Defined	Defined
Plant Location	General	Approximate	Specific	Specific	Specific
Soils & Hydrology	None	Preliminary	Defined	Defined	Defined
Integrated Project Plan	None	Preliminary	Defined	Defined	Defined
Project Master Schedule	None	Preliminary	Defined	Defined	Defined
Escalation Strategy	None	Preliminary	Defined	Defined	Defined
Work Breakdown Structure	None	Preliminary	Defined	Defined	Defined
Project Code of Accounts	None	Preliminary	Defined	Defined	Defined
Contracting Strategy	Assumed	Assumed	Preliminary	Defined	Defined
Engineering Deliverables:					
Block Flow Diagrams	S/P	P/C	C	C	C
Plot Plans		S/P	C	C	C
Process Flow Diagrams (PFDs)		P	C	C	C
Utility Flow Diagrams (UFDs)		S/P	C	C	C
Piping & Instrument Diagrams (P&IDs)		S/P	C	C	C
Heat & Material Balances		S/P	C	C	C
Process Equipment List		S/P	C	C	C
Utility Equipment List		S/P	C	C	C
Electrical One-Line Drawings		S/P	C	C	C
Specifications & Datasheets		S	P/C	C	C
General Equipment Arrangement Drawings		S	C	C	C
Spare Parts Listings			P	P	C
Mechanical Discipline Drawings			S/P	P/C	C
Electrical Discipline Drawings			S/P	P/C	C
Instrumentation/Control System Discipline Drawings			S/P	P/C	C
Civil/Structural/Site Discipline Drawings			S/P	P/C	C

Table 3 – Estimate Input Checklist and Maturity Matrix (Primary Classification Determinate)

REFERENCES

1. AACE International, Recommended Practice No.17R-97, *Cost Estimate Classification System*, AACE International, Morgantown, WV. (latest revision)
2. Hollmann, John K., PE CCE, Editor, *Total Cost Management Framework: An Integrated Approach to Portfolio, Program and Project Management*, AACE International, Morgantown, WV, 2006.
3. AACE International, Recommended Practice 10S-90, *Cost Engineering Terminology*, AACE International, Morgantown, WV. (latest revision).
4. John R. Heizelman, *Estimating Factors for Process Plants*, 1988 AACE Transactions, V.3, AACE International, Morgantown, WV, 1988.
5. K.T. Yeo, The Cost Engineer Journal, UK Vol. 27, No. 6, 1989.
6. Stevens, G. and T. Davis, How Accurate are Capital Cost Estimates?, 1988 AACE Transactions, B.4, AACE International. Morgantown, WV, 1988. (* Class 3 is inferred)
7. Behrenbruch, Peter, article in Journal of Petroleum Technology, Vol. 45, No. 8, Society of Petroleum Engineers, August 1993.
8. AACE International, Recommended Practice 42R-08, *Risk Analysis and Contingency Determination Using Parametric Estimating*, AACE International, Morgantown, WV, (latest revision).

CONTRIBUTORS

Peter Christensen, CCE (Author)
Larry R. Dysert, CCC CEP (Author)
Jennifer Bates, CCE
Jeffery J. Borowicz, CCE CEP PSP
Peter R. Bredehoeft, Jr. CEP
Robert B. Brown, PE
Dorothy J. Burton
Robert C. Creese, PE CCE
John K. Hollmann, PE CCE CEP
Kenneth K. Humphreys, PE CCE
Donald F. McDonald, Jr. PE CCE PSP
C. Arthur Miller
Todd W. Pickett, CCC CEP
Bernard A. Pietlock, CCC CEP
Wesley R. Querns, CCE
Don L. Short, II CEP
H. Lance Stephenson, CCC
James D. Whiteside, II PE