Megaproject policy and planning
problems, causes, cures
Flyvbjerg, Bent

Publication date:
2007

Document Version
Også kaldet Forlagets PDF

Link to publication from Aalborg University

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

? Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
? You may not further distribute the material or use it for any profit-making activity or commercial gain
? You may freely distribute the URL identifying the publication in the public portal?

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.
Indholdsfortegnelse/Contents

Dansk resumé (in Danish) ................................................................. v
Problemer i policy og planlægning for megaprojekter ......................... v
Forklaringsmodeller ........................................................................ vi
Løsninger ....................................................................................... vii
Konklusion ................................................................................... viii

Sammenfattende redegørelse, Summary Paper (in English) .............. 1
Abstract ....................................................................................... 2
Introduction .................................................................................. 2

Methodology .................................................................................. 3
  Four Steps to Understanding Misleading Estimates of Costs,
  Benefits, and Risks .................................................................. 3
  Overall Design of the Research .............................................. 4
  Sampling and Data Collection ............................................... 5
  In-Depth Interviews ............................................................... 10

Problems in Megaproject Policy and Planning ................................. 12
  The Size of Cost Overruns and Benefit Shortfalls .................. 13
  Examples of Cost Overruns and Benefit Shortfalls .............. 14
  Why Cost Overruns and Benefit Shortfalls Are a Problem ........ 16
  Policy Implications ............................................................... 17

Causes of Inaccurate Costs and Benefits ....................................... 18
  Three Explanations of Cost Overruns and Benefit Shortfalls ...... 18
  Validity of Explanations ....................................................... 19
  Towards Falsifiable Theories .................................................. 25

Cures: The Need for Reform in Megaproject Policy and Planning ....... 27
  Better Methods: Reference Class Forecasting ....................... 28
  Improved Incentives: Public and Private Sector Accountability .... 32
Dansk resumé af doktorafhandlingen "Megaprocess Policy and Planning: Problems, Causes, Cures"

Megaprojektets politik og planlægning: Problemer, årsager, løsninger

Af Bent Flyvbjerg

Megaprojekter defineres som de dyreste infrastruktur- og investeringsprojekter, der udføres i verden i dag, typisk til priser pr. projekt fra nogle hundrede millioner til milliarder af dollars. Afhandlingen fokuserer på tre hovedforhold ved policy og planlægning for megaprojekter: (1) Problemer i policy og planlægning for megaprojekter, (2) forklaringsmodeller for de identificerede problemer, og (3) metoder til løsning af problemerne, såvel tekniske som institutionelt-organisatoriske.

Afhandlingens hovedfokus er på store transportinfrastrukturprojekter, men ved hjælp af sammenlignende studier vises det, at afhandlingens konklusioner også gælder andre projekttyper, f.eks. dæmninger, kraftværker, vandprojekter, sports- og kulturbygninger, messe- og konferencecentre, IT systemer, olie- og gasudvindingsprojekter, rumforskningprojekter og våbensystemer. Afhandlingens metoder er statistisk og komparativ analyse, dybdebørendte interviews, spørgeskemaundersøgelser og teoretiske studier.

Problemer i policy og planlægning for megaprojekter

Afhandlingen identificerer misinformation om projekters omkostninger, benefits og risici som det vigtigste problem i policy og planlægning for megaprojekter. En konsekvens af denne misinformation er budgetoverskridelser, manglende indtægter og ressourcespild.

For transportinfrastrukturprojekter gælder det, at 9 ud af 10 projekter har overskridelser i anlægsbudgettet. Den gennemsnitlige overskridelse for baner er 44,7%, for broer og tunneler 33,8% og for veje 20,4%, målt i faste priser og konservativt vurderet. Hertil kommer, at standardafvigelserne på de gennemsnitlige overskridelser er store, hvilket samlet viser, at den økonomiske risiko på anlægssiden er særdeles høj.
På benefit-siden gælder det, at 9 ud af 10 baneprojekter har overvurderede passagerprognoser. Den gennemsnitlige overvurdering er her 105,6%, hvilket svarer til, at mere end halvdelen af de prognosticerede passagerer aldrig dukker op. 84% af passagerprognoserne er mere end 20% forkerte. For veje er 50% af trafikprognoserne mere end 20% forkerte, men med en mere lige fordeling mellem over- og undervurderet trafik end for baner. Også på benefit-siden er standardafvigelserne store, og igen er dette identisk med høj økonomisk risiko.

En konsekvens af de misvisende estimator af omkostninger og benefits er, at de cost-benefit analyser, som typisk bruges til at begrunde store infrastrukturinvesteringer, bliver stærkt misvisende. Ikke blot med nogle få procent, men ofte med flere faktorer. Det samme gælder socio-økonomiske vurderinger og vurderinger af virkninger for miljøet.

Afhandlingen viser, at problemet med misinformation om omkostninger, benefits og risici er konstant over rum og tid. Fænomenet findes i samtlige undersøgte lande, d.v.s. 20 lande på 5 kontinenter for omkostninger og 14 lande på 5 kontinenter for benefits. Budgetoverskridelserne har ligget på samme høje niveau og trafikprognosernes nøjagtighed er ikke forbedret for den periode, som afhandlingen dækker, hvilket vil sige 70 år for omkostninger og 30 år for trafik. Hvis metoderne til at prognosticere omkostninger og efterspørgsel er forbedret over tid, så er det ikke noget, som kan spores i de empiriske data.

**Forklaringsmodeller**

Afhandlingen beskriver og afprøver tre hovedmodeller til forklaring af problemet med misinformation: Tekniske modeller (fejl i data og metoder), psykologiske modeller ("the planning fallacy", "optimism bias") og politisk-økonomiske modeller (strategisk misrepræsentation, løgn).

Afhandlingen forkaster tekniske og psykologiske forklaringsmodeller på grundlag af en vurdering af deres "fit" med de statistiske data samt resultater fra spørgeskema- og interviewundersøgelser. Dette forstås i afhandlingen ikke således, at der ikke findes situationer, hvor de to modeller har forklaringskraft. Men disse situationer er ikke tilstrækkelig typiske for planlægning af megaprojekter, til at modellerne opnår en acceptabel forklaringsgrad her.
For planlægning af megaprojekter viser afhandlingen, at politisk-økonomiske forklaringsmodeller passer bedst med de eksisterende data, og det statistiske "fit" bekræftes substantivt gennem interviews og spørgeskemaundersøgelser med hovedaktører involveret i policy og planlægning for megaprojekter. Planlæggere og projektmagere misinformerer således aktivt og med vilje om de omkostninger, benefits og risici, som er forbundet med deres projekter, med henblik på at øge sandsynligheden for, at det er netop deres projekt, og ikke konkurrenternes, som bliver godkendt og finansieret.


Løsninger

Endelig præsenterer afhandlingen en række reform-foranstaltninger, som er møntet på at komme misinformation og den omvendte Darwinisme til livs, og dermed forbedre policy og planlægning for megaprojekter. Foranstaltningerne dækker to hovedområder: (1) Udvikling af bedre metoder til at estimere omkostninger, benefits og risici, og (2) etablering af incitament-strukturer, der vil begrænse misinformation og fremme information. Af de to typer foranstaltninger vurderes ændrede incitament-strukturer at være vigtigst.


Til forbedring af incitament-strukturen for policy og planlægning af megaprojekter argumenterer afhandlingen for en pakke af tiltag omfattende bl.a. bloktilskud til infrastruktur i stedet for øremærkede midler; uafhængige peer reviews og benchmarking af alle prognoser; større offentlig indsigt i prognoser; brug af borgerjuryer o.l. til vurdering af projekter og deres prognoser; vurdering af prognoser på faglige konference; aktiv belønning af nøjagtige prognoser; professionel og juridisk straf til personer og organisationer, som bevidst producerer vildledende prognoser; afskaffelse af fuld statsgaranti til projekter; minimum en tredjedel privat finansiering af projekter som forudsætning for projektstart, også for subsidierede projekter; kontrakter, som gør prognosemagere og deres organisationer finansielt medansvarlige for budgetoverskridelser og manglende benefits; og endelig én og kun én ansvarlig projektorganisation for et projekt.

Hvis incitament-strukturen ændres på denne eller lignende måde, vil det ifølge afhandlingen være muligt at begrænse misinformationen om megaprojekter og nå frem til bedre projekter. Hvis incitament-strukturen ikke ændres vil misinformationen sandsynligvis fortsætte og brugen af midler til megaprojekter vil fortsætte med at være ineffektiv og udemokratisk.

**Konklusion**

Afslutningsvis identificerer afhandlingen to tendenser, som for tiden driver policy og planlægning af megaprojekter i den rigtige retning, d.v.s. mod større effektivitet og mere demokrati. For det første har principperne for "good governance" vokset sig stærkere i de seneste år i kølvandet på Enron skandalen og dens mange efterfølgere. Disse principper har også fundet vej til den offentlige sektor og med samme formål, nemlig at bremse finansielt misbrug og fremme bedre administration. Selvom udviklingen er langsom, mærkes den også i policy og planlægning for megaprojekter.
For det andet har megaprojekter efterhånden vokset sig så store i forhold til nationale økonomier, at budgetoverskridelser, manglende efterspørgsel, og risici fra bare et enkelt projekt kan destabilisere finanserne i et helt land eller en hel region, som det er sket i f.eks. Grækenland og Hong Kong. Parlaments og regeringer er begyndt at indse, at nationaløkonomiske problemer er for stor en pris at betale for den konventionelle måde at planlægge megaprojekter. Storbritannien og Holland er foregangslande i denne sammenhæng og kan studeres som "best practice".

På den baggrund konkluderer afhandlingen, at der er grundlag for en vis optimisme med hensyn til planlægning og policy for megaprojekter i de kommende år.
Sammenfattende redegørelse for doktorafhandlingen
Megaproject Policy and Planning: Problems, Causes, Cures

Af

Bent Flyvbjerg
Aalborg Universitet

Version 2.0
October, 2006
Abstract
This paper focuses on problems in megaproject policy and planning and their causes and possible cures. After considerations of methodology, the paper first identifies as a main problem in megaproject development pervasive misinformation about the costs, benefits, and risks involved. A consequence of misinformation is cost overruns, benefit shortfalls, and waste. Second, the paper explores the causes of misinformation and finds that political-economic explanations best account for the available evidence: planners and promoters deliberately misrepresent costs, benefits, and risks in order to increase the likelihood that it is their projects, and not the competition's, that gain approval and funding. This results in the "survival of the unfittest," where often it is not the best projects that are built, but the most misrepresented ones. Finally, the paper presents measures for reforming megaproject policy and planning with a focus on better planning methods and improved governance structures, the latter being more important.

Introduction
For a number of years my research group and I have explored different aspects of megaproject policy and planning (Flyvbjerg, Bruzelius, and Rothengatter, 2003; Flyvbjerg, Holm, and Buhl, 2002, 2004, 2005; Flyvbjerg and Cowi, 2004; Flyvbjerg, 2005a, 2005b). In this paper I would like to take stock of what we have learned from our research so far.

First I will argue that a major problem in megaproject policy and planning is the high level of misinformation about costs and benefits that decision makers face in deciding whether to build, and the high risks such misinformation generates. Second I will explore the causes of misinformation and risk, mainly in the guise of optimism bias and strategic misrepresentation. Finally, I will present a number of measures aimed at better policy and planning, including changed incentive structures and better planning methods. Thus the paper is organized as a simple triptych consisting in problems, causes, and cures.

The emphasis will be on transportation infrastructure projects. It should be mentioned at the outset, however, that comparative research shows that the problems, causes, and cures identified for transportation apply to a wide range of other project types including power plants, dams, water projects, concert halls, museums, sports arenas, convention centers, IT systems, oil and gas extraction projects, aerospace
projects, and weapons systems (Flyvbjerg, Bruzelius, and Rothengatter, 2003: 18-19; Flyvbjerg, Holm, and Buhl, 2002: 286; Flyvbjerg, 2005a; Altshuler and Luberoff, 2003).

**Methodology**

This section focuses on key methodological questions of the research summarized in the present paper. First, the research is placed in the context of the overall body of research on misinformation in estimates of costs, benefits, and risks in transportation infrastructure policy and planning. Second, five key steps are identified in the methodological design of the research. Third, issues of sampling and data collection are discussed. Finally, questions of interviewing are discussed.

*Four Steps to Understanding Misleading Estimates of Costs, Benefits, and Risks*

Four steps may be identified in the evolution of a body of scholarly research aimed at understanding misinformation in estimates of costs, benefits, and risks in large-scale transportation investments. The first step was taken by Pickrell (1990) and Fouracre, Allport, and Thomson (1990), who provided sound evidence for a small number of urban rail projects that substantial misinformation about costs and benefits is a problem, and who implied that such misinformation may be caused by deception on the part of project promoters and forecasters.

The second step was taken by Wachs (1990), who established--again for a small sample of urban rail projects--that lying, understood as intentional deception, is, in fact, an important cause of misinformation about costs, benefits, and risks. Wachs began the difficult task of charting who does the lying--or "strategic misrepresentation" as infrastructure professionals like to call it--why it occurs, what the ethical implications are, etc. Wachs's conclusions for the US were later supported by a study for the UK carried out by Flyvbjerg and Cowi (2004).

The problem with the research in the first two steps is that it is based on too few cases to be statistically significant; the pattern found may be due to random properties of the small samples involved. This problem is solved in the third step, taken with the work summarized in the present paper. Based on a large sample of transportation infrastructure projects, it is shown that (1) the pattern of misinformation uncovered by Pickrell and others is of general import and is statistically significant, and (2) the pattern is robust
across different project types, different geographical regions, and different historical periods. It is also shown that the large sample presented here lends statistical support to the conclusions about strategic misrepresentation arrived at by Wachs for his small sample.

The fourth and final step in understanding misinformation would be to do for a large sample of different transportation infrastructure projects what Wachs (1990) and Flyvbjerg and Cowi (2004) did for smaller samples: establish the motives and intentions of forecasters and promoters and how they affect misinformation. This may be done by surveying with questionnaires and interviews a large number of forecasters and project promoters, representing a large number of projects, and asking them about what happened in specific instances of misleading estimates of costs and benefits, why estimates turned out wrong, what the forecasters’ reasons were for making the estimates as they did, etc. Brinkman (2003) has taken a first step in this direction, but much interesting work remains to be done.

Overall Design of the Research
The design of the research summarized in the present paper may be described by the following main steps:

1. First, a large sample of transportation infrastructure projects was established with the objective to falsify the thesis that cost overruns and benefit shortfalls are common in such projects. The thesis could not be falsified. At a very high level of statistical significance it was established that cost overruns and benefit shortfalls are common in transportation infrastructure development. Data on the size and frequency of cost overruns and benefit shortfalls was collected by means of questionnaires, from annual reports, from available government statistics, national auditors, other studies, etc. Statistical and comparative methods were used to analyze variation in the data across project types, nations, continents, and time.

2. Second, theoretical studies were carried out in order to identify explanatory models for cost overruns and benefit shortfalls. Four types of explanatory models were identified: Technical, economic, psychological, and political.
3. Third, statistical analysis and in-depth interviews were performed to test the validity of explanatory models for cost overruns and benefit shortfalls. It was found that political and psychological explanations best accounted for the data.

4. Fourth, based on theories that validly explain cost overruns and benefit shortfalls, measures were developed that may help curb overruns and shortfalls. Two types of measures were identified: (a) Better methods for estimating costs and benefits, especially reference class forecasting, and (b) Measures of better governance, especially improved accountability.

5. Finally, through comparative studies it was established to what extent the results found for transportation infrastructure may be expected to be valid for other project types. The available evidence suggests that cost overruns and benefit shortfalls are neither larger nor smaller for transportation infrastructure than for other project types, including power plants, dams, water projects, concert halls, museums, sports arenas, convention centers, IT systems, oil and gas extraction projects, aerospace projects, and weapons systems.

The following two sections focus on basic methodological questions of how the sample of transportation infrastructure projects was established and how the persons for in-depth interviews were chosen.

**Sampling and Data Collection**

The first task of the research reported in this paper was to establish a sample of infrastructure projects substantially larger than what is common in this area of research. The sample had to be large enough to allow statistical analyses of cost overruns and benefit shortfalls. Here a first problem was that data on actual costs and benefits in transportation infrastructure projects are relatively scarce. One reason is that it is quite time consuming to unearth such data.

For cost data, in public sector projects funding and accounting procedures are typically unfit for keeping track of the multiple and complex changes that occur in total project costs over time. For large projects, the relevant time frame may cover 5, 10, or more fiscal years from planning, until decision to build, until construction starts, until the project is completed and operations begin. Reconstructing the actual total costs of a
public project, therefore, typically entails long and difficult archival work and complex accounting. For private projects, even if funding and accounting practices may be more conducive to producing data on actual total costs, such data are often classified to keep them from the hands of competitors. Unfortunately, this also tends to keep data from the hands of scholars. And for both public and private projects, data on actual costs may be held back by project owners because more often than not, actual costs reveal substantial cost overrun, and overrun is normally considered somewhat of an embarrassment to promoters and owners. In sum, establishing reliable data on actual costs for even a single transportation infrastructure project is often highly time consuming or simply impossible.

For data on travel demand, in public sector projects it is surprising how often count data on actual traffic are simply not produced. And even where the intention is to produce the data, projects may develop in ways that make it difficult or impossible to compare forecasted with actual traffic. For example, forecasted traffic for a given project may be estimated for the opening year, but due to delays, which are common, the actual opening date turns out to be several years later than that forecasted, and no forecast of traffic was made for the actual opening year. In more general terms, methodological differences in how and on what basis forecasted and actual traffic are estimated often make comparisons difficult or impossible. Finally, for large projects the elapse of time from forecasts are made, until actual traffic can be counted may cover 10-15 years. Over such long periods, the assumptions underlying forecasts may be dated and incommensurate when compared with the assumption underlying the way actual traffic is measured; or initial plans to compare actual with forecasted traffic may be given up or simply forgotten. For private sector projects, traffic typically generates an income for the project owner. Budgeting and accounting is commercial, and therefore traffic forecasts and traffic counts tend to be more systematic and more conducive to comparative studies of forecasted and actual traffic than is the case for public sector projects. This typically does not help scholars much, however, because again data are often classified to keep them from competitors. And for both public and private projects, data that allow forecasted and actual traffic to be compared may be held back by project owners and managers because the size and direction of differences between forecasted and actual traffic may be of a kind that, if made public, would make the project look bad in the public eye, for instance where actual traffic is substantially lower than that forecasted.

This state of affairs explains why small-sample studies dominate scholarship in this field of research. But despite the problems mentioned, after several years of data collection and refinement, for costs it was possible to establish a sample of 258 trans-
portation infrastructure projects with data on both actual construction costs and estimated costs at the time of decision to build. The project portfolio is worth approximately US$90 billion (1995 prices). The project types are bridges, tunnels, highways, freeways, high-speed rail, urban rail, and conventional (interurban) rail. The projects are located in 20 countries on 5 continents, including both developed and developing nations. The projects were completed between 1927 and 1998. Older projects were included in the sample in order to test whether the accuracy of estimated costs improved over time. The construction costs of projects range from US$1.5 million to US$8.5 billion (1995 prices), with the smallest projects typically being stretches of roads in larger road schemes, and the largest projects being rail links, tunnels, and bridges.

For traffic, it was possible to establish comparable data for forecasted and actual traffic for 210 projects located in 14 countries on five continents, again including both developed and developing nations. As far as is known, this is the largest sample of projects with comparable data on forecasted and actual costs and traffic that has been established for this type of project.

In statistical analysis, data should be a sample from a larger population, and the sample should represent the population properly. These requirements are ideally satisfied by drawing the sample by randomized lot. Randomization ensures with high probability that factors that cannot be controlled are equalized. A sample should also be designed such that the representation of subgroups corresponds to their occurrence and importance in the population. In studies of human affairs, however, where controlled laboratory experiments often cannot be conducted, it is frequently impossible to meet these ideal conditions. This is also the case for the current study, and a different approach to sampling and statistical analysis therefore had to be taken.

The projects for the sample were selected on the basis of data availability. All projects that the research team knew of for which data on cost and/or traffic development were obtainable were considered for inclusion in the sample. Cost development is defined as the difference between actual and estimated costs in percentage of estimated costs, with all costs measured in fixed prices. Traffic development is similarly defined as the difference between actual and estimated traffic in percentage of estimated traffic.

Data on cost development defined in this manner were available for 343 projects. 85 projects were then rejected because of insufficient data quality. For instance, for some projects it was not possible to obtain a clear answer regarding what was included in costs, or whether cost data were given in current or fixed prices, or which price level (year) had been used in estimating and discounting costs. More specifically, of those 85
projects, 27 were rejected because it could not be established whether or not cost data were valid and reliable. 12 projects were rejected because they had been completed before 1915 and no reliable indices were available for discounting costs to the present. Finally, 46 projects were excluded because cost development for them turned out to have been calculated before construction was completed and operations begun; therefore, the actual final costs for these projects may be different from the cost estimates used to calculate cost development, and no information was available on actual final costs. In addition to the 85 rejected projects mentioned here, a number of projects were also rejected to avoid double counting of projects. This typically involved projects from other studies that appeared in more than one study or where there was a strong suspicion that this might be the case.

Data on traffic development were available for 485 projects. A total of 275 projects were then rejected because of unclear or insufficient data quality. Of the 275 projects rejected, 124 were rejected because inaccuracy had been estimated in ways different from and incomparable to the way it was decided to estimate inaccuracy; 151 projects were rejected because inaccuracies for these projects had been estimated on the basis of adjusted data for actual traffic instead of using original, actual data as it was decided to do. In sum, all projects for which valid and reliable data were available for cost and traffic development were included in the sample.

As for any sample, a key question is whether the sample is representative of the population. Here the question is whether the projects included in the sample are representative of the population of transportation infrastructure projects. Since the criterion for sampling was data availability, this question translates into one of whether projects with available data are representative. There are four reasons why this is probably not the case. First, it may be speculated that projects that are managed well with respect to data availability may also be managed well in other respects, resulting in better than average (i.e., nonrepresentative) performance for such projects. Second, it has been argued that the very existence of data that make the evaluation of performance possible may contribute to improved performance when such data are used by project management to monitor projects (World Bank, 1994, p. 17). Again, such projects would not be representative of the project population. Third, one may speculate that managers of projects with a particularly bad track record regarding cost overruns and/or traffic shortfalls have an interest in not making data available, which would then result in underrepresentation of such projects in the sample. Conversely, managers of projects with a good track record for costs and traffic might be interested in making this public, resulting in overrepresen-
tation of these projects. Fourth, and finally, even where managers have made cost and traffic data available, they may have chosen to give out data that present their projects in as favorable a light as possible. Often there are several estimates of costs and traffic to choose from and several calculations of actual costs and traffic for a given project at a given time. If researchers collect data by means of survey questionnaires, as is often the case, there might be a temptation for managers to choose the combination of actual and estimated costs and traffic that suits them best, possibly a combination that makes their projects look good.

The available data do not allow an exact, empirical assessment of the magnitude of the problem of representativeness. But it is concluded, for the reasons given above, that most likely the sample is biased and the bias is conservative. In other words, cost overruns and traffic shortfalls estimated from the sample would be lower than those in the project population. This should be kept in mind when interpreting the results from the statistical analyses below. The sample of projects is not perfect by any means. Still, it is the best obtainable sample given the current state-of-the-art in this field of research.

In the statistical analyses, percentage cost and traffic development in the sample is considered normally distributed unless otherwise stated. Residual plots, not shown here, indicate that normal distribution might not be completely satisfied. However, transformations (e.g., the logarithmic one) do not improve this significantly. For simplicity, therefore, no transformation has been made, unless otherwise stated. The subdivisions of the sample implemented as part of analyses entail methodological problems of their own. Thus the representation of observations in different combinations of subgroups is quite skewed for the data considered. The analysis would be improved considerably if the representation were more even. Partial and complete confounding occur; that is, if a combination of two or more effects is significant, it is sometimes difficult to decide whether one, the other, or both cause the difference. For interactions, often not all the combinations are represented, or the representations can be quite scarce. The interpretations of the data have been adapted to these limitations, needless to say. If better data could be gathered, sharper conclusions could be made.

The statistical models used are linear normal models, i.e., analysis of variance and regression analysis with the appropriate F-tests and t-tests have been made. The tests of hypotheses concerning mean values are known to be robust to deviations from normality. Also, chi-square tests for independence have been used for count data. For each test, the p-value has been reported.
Finally, it should be mentioned that although the sample of transportation infrastructure projects used in this study is the largest of its kind, it is still too small to allow more than a few subdivisions, if comparative statistical analyses must still be possible. Therefore, in further work on understanding misinformation in estimates of costs, benefits, and risks, the sample should be enlarged to better represent different types of projects and different geographical locations. As to project types, data for more private projects would be particularly useful in allowing statistically valid comparisons between public and private sector projects. Such comparisons do not exist today, and nobody knows whether private projects perform better or worse than public ones regarding misinformation. The sample should also be enlarged to contain data for more fixed links and rail projects. Such data would allow a better (i.e., a statistically corroborated) comparative understanding of misinformation for more specific subtypes of projects like bridges, tunnels, high-speed rail, urban rail, and conventional rail. Such an understanding is non-existent today. As to geography, immediate rewards would be gained from data for projects outside Europe and North America, especially for fixed links and roads. But even for Europe and North America, data on more projects are needed to allow more detailed comparative analysis.

In-Depth Interviews
A key question for the body of research summarized in this paper is whether actors deliberately produce the cost underestimates and benefit overestimates documented by the research, or whether such under- and overestimates are non-intentional errors.

This question raises the difficult issue of deception and lying. Questions of lying are notoriously hard to answer, because per definition a lie is making a statement intended to deceive others, and in order to establish whether lying has taken place, one must therefore know the intentions of actors. In research, intentions are typically uncovered by interviews. But for legal, economic, moral, and other reasons, if actors have intentionally manipulated estimates of costs and benefits to get a project started, they are unlikely to formally tell researchers or others that this is the case.

To illustrate just how difficult it is to obtain this type of information, consider the approval to build the multibillion-dollar Øresund bridge between Denmark and Sweden. After the Danish Parliament approved building the bridge, it appeared that officials in the Danish Ministry of Transport had concealed from Parliament several cost-benefit analyses that showed the bridge would be non-viable. Instead, the ministry had provided
Parliament with estimates of costs and benefits that seemed to have been manipulated to show viability. Later, when this state of affairs was confirmed by the National Audit Office of Denmark (1993), the result was a sharp criticism of the Ministry of Transport. It is unlawful to misinform Parliament, and for a while there was talk of conducting a formal investigation to place responsibility for the misdemeanor.

One of my former students was involved, as a civil servant, in producing the estimates of costs and benefits that were given to parliament. When we met at a conference, I suggested I interview him about how the estimates were arrived at. I argued that the Øresund bridge is a historically important project, linking Sweden with Denmark, and Scandinavia with Europe, for the first time. The bridge is also one of the largest cross-national public works projects in the world, plus among the two or three biggest public investments in Danish history. For these reasons, posterity would clearly have an interest in knowing how the decision to build the bridge was arrived at, or so I argued.

I knew the issue was sensitive and I therefore proposed to my former student that he could decide such terms for the uses of the interview that he would be comfortable with. He could decide, for instance, that the interview be placed in a sealed envelope with the Danish National Archives, not to be opened until a date decided by him, fifty or a hundred years in the future if so he preferred. Furthermore, I stated my willingness to sign as interviewer a legally binding statement of nondisclosure and nonuse of the interview.

With these conditions, still my former student was visibly uncomfortable about even discussing the prospect of an interview. He said the interview was not possible, excused himself, and hurried away. He appeared profoundly scared, perhaps because he knew that any acknowledgment on his part that he or his colleagues had deliberately misinformed the minister of the costs and benefits of the Øresund bridge, or that the minister had deliberately misinformed Parliament, would stamp them as lawbreakers.

It is that difficult to get those who work in environments with political and organizational pressure to talk about what really goes on when costs are underestimated and benefits overestimated in decision making about large infrastructure projects. This makes deception and lying unusually difficult topics for empirical, scholarly research. As a consequence, only few studies exist.

As far as is known, only two studies exist that have succeeded in getting those involved in underestimating costs and overestimating benefits in transportation infrastructure projects to talk about issues of deliberation and intention. The first study was carried out by Flyvbjerg and Cowi (2004) and covered interviews with 13 public
officials, planners, and consultants working in UK transportation policy and planning as developers, approvers, or implementers of projects. The interviewees worked with national government, local transport providers, and private consultancies. The interviewees were chosen for their detailed knowledge of the project development process and with a view to covering both central and local government and both the public and private sectors. The study was endorsed by the UK Department for Transport, which made interviewees more forthcoming with researchers than is usually the case.

The second study was carried out by Wachs (1990). This study, which Wachs describes as "case study research" covered conversations with approximately 15 planners, managers, and consultants. The conversations were not structured interviews and did not use a standard instrument or survey design (personal communication with Martin Wachs, 2006, author's archives). Wachs's study may be described as based on exploratory, unstructured interviews, the methodology reflecting the early, pioneering stage of the field.

Ideally, it would be desirable to interview more public officials, planners, and consultants than done by Flyvbjerg and Cowi (2004) and Wachs (1990). Establishing a larger sample of actors willing to be interviewed is a key challenge for future research, and a challenge that realistically will be particularly difficult to meet for the reasons presented above. At present, the data from the interviews carried out by Flyvbjerg and Cowi (2004) and Wachs (1990) are the best that exist regarding whether actors deliberately underestimate costs and overestimate benefits as part of the project development process. The data do not present final proof for the role of deliberation in forecasting. They do, however, constitute one significant step in a cumulative research process for testing whether deliberation plays a role in forecasting and to what extent.

**Problems in Megaproject Policy and Planning**

Megaprojects generally have the following characteristics:

- Such projects are inherently risky due to long planning horizons and complex interfaces.
- Decision making, policy, and planning is often multi-actor processes with conflicting interests.
- Often the project scope or ambition level changes significantly over time.
Statistical evidence shows that such unplanned events are often unaccounted for, leaving budget and other contingencies sorely inadequate. As a consequence, misinformation about costs, benefits, and risks is the norm. The result is cost overruns and/or benefit shortfalls with a majority of projects.

**The Size of Cost Overruns and Benefit Shortfalls**

For transportation infrastructure projects, Table 1 shows the inaccuracy of construction cost estimates measured as the size of cost overrun. For rail, average cost overrun is 44.7 percent measured in constant prices. For bridges and tunnels, the equivalent figure is 33.8 percent, and for roads 20.4 percent. The difference in cost overrun between the three project types is statistically significant, indicating that each type should be treated separately (Flyvbjerg, Holm, and Buhl, 2002).

The large standard deviations shown in Table 1 are as interesting as the large average cost overruns. The size of the standard deviations demonstrate that uncertainty and risk regarding cost overruns are large, indeed.

The following key observations pertain to cost overruns in transportation infrastructure projects:

- 9 out of 10 projects have cost overrun.
- Overrun is found in the 20 nations and 5 continents covered by the study.
- Overrun is constant for the 70-year period covered by the study, estimates have not improved over time.

Table 2 shows the inaccuracy of travel demand forecasts for rail and road projects. For rail, actual passenger traffic is 51.4 percent lower than estimated traffic on average. This is equivalent to an average overestimate in rail passenger forecasts of no less than 105.6 percent. The result is large benefit shortfalls for rail. For roads, actual vehicle traffic is on average 9.5 percent higher than forecasted traffic. We see that rail passenger forecasts are biased, whereas this is not the case for road traffic forecasts. The difference between rail and road is statistically significant at a high level. Again the standard deviations are large, indicating that forecasting errors vary widely across projects (Flyvbjerg, Holm, and Buhl, 2005; Flyvbjerg, 2005b).
The following observations hold for traffic demand forecasts:

- 84 percent of rail passenger forecasts are wrong by more than ±20 percent.
- 9 out of 10 rail projects have overestimated traffic.
- 50 percent of road traffic forecasts are wrong by more than ±20 percent.
- The number of roads with overestimated and underestimated traffic, respectively, is about the same.
- Inaccuracy in traffic forecasts are found in the 14 nations and 5 continents covered by the study.
- Inaccuracy is constant for the 30-year period covered by the study, forecasts have not improved over time.

It is concluded that if techniques and skills for arriving at accurate cost and traffic forecasts have improved over time, these improvements have not resulted in an increase in the accuracy of forecasts.

[Table 2 app. here]

If we combine the data in tables 1 and 2, we see that for rail an average cost overrun of 44.7 percent combines with an average traffic shortfall of 51.4 percent. For roads, an average cost overrun of 20.4 percent combines with a fifty-fifty chance that traffic is also wrong by more than 20 percent. As a consequence, cost benefit analyses and social and environmental impact assessments based on cost and traffic forecasts like those described above will typically be highly misleading.

**Examples of Cost Overruns and Benefit Shortfalls**

The list of examples of projects with cost overruns and/or benefit shortfalls is seemingly endless (Flyvbjerg, 2005a). Boston’s Big Dig, otherwise known as the Central Artery/Tunnel Project, were 275 percent or US$11 billion over budget in constant dollars when it opened, and further overruns are accruing due to faulty construction. Actual costs for Denver’s $5 billion International Airport were close to 200 percent higher than estimated costs. The overrun on the San Francisco-Oakland Bay Bridge retrofit was $2.5 billion, or more than 100 percent, even before construction started. The Copenhagen metro and many other urban rail projects worldwide have had similar overruns. The
Channel tunnel between the UK and France came in 80 percent over budget for construction and 140 percent over for financing. At the initial public offering, Eurotunnel, the private owner of the tunnel, lured investors by telling them that 10 percent “would be a reasonable allowance for the possible impact of unforeseen circumstances on construction costs.” Outside of transportation, the $4 billion cost overrun for the Pentagon spy satellite program and the over $5 billion overrun on the International Space Station are typical of defense and aerospace projects. Studies show that large infrastructure and technology projects tend statistically to follow a pattern of cost underestimation and overrun. Many such projects end up financial disasters. Unfortunately, the consequences are not always only financial, as is illustrated by the NASA space shuttle. Here, the cooking of budgets to make this under-performing project look good on paper has been linked with shortchanged safety upgrades related to the deaths of seven astronauts aboard the Columbia shuttle in 2003 (Flyvbjerg 2004).

As for benefit shortfalls, consider Bangkok’s US$2 billion Skytrain, a two-track elevated urban rail system designed to service some of the most densely populated areas from the air. The system is greatly oversized, with station platforms too long for its shortened trains. Many trains and cars sit in the garage, because there is no need for them. Terminals are too large, etc. The reason is that actual traffic turned out to be less than half that forecast (Flyvbjerg, Holm, and Buhl, 2005: 132). Every effort has been made to market and promote the train, but the project company has ended up in financial trouble. Even though urban rail is probably a good idea for a dense, congested, and air-polluted city like Bangkok, overinvesting in idle capacity is hardly the best way to use resources, especially in a developing nation in which capital for investment is scarce. Such benefit shortfalls are common and have also haunted the Channel tunnel, the Los Angeles and Copenhagen metros, and Denver’s International Airport.

Other projects with cost overruns and/or benefit shortfalls are, in North America: the F/A-22 fighter aircraft; FBI’s Trilogy information system; Ontario’s Pickering nuclear plant; subways in numerous cities, including Miami and Mexico City; convention centers in Houston, Los Angeles, and other cities; the Animas-La Plata water project; the Sacramento regional sewer system renewal; the Quebec Olympic stadium; Toronto’s Sky Dome; the Washington Public Power Supply System; and the Iraq reconstruction effort. In Europe: the Eurofighter military jet, the new British Library, the Millennium Dome, the Nimrod maritime patrol plane, the UK West Coast rail upgrade and the related Railtrack fiscal collapse, the Astute attack submarine, the Humber Bridge, the Tyne metro system, the Scottish parliament building, the French Paris Nord
TGV, the Berlin-Hamburg maglev train, Hanover's Expo 2000, Athens' 2004 Olympics, Russia's Sakhalin-1 oil and gas project, Norway's Gardermoen airport, the Øresund Bridge between Sweden and Denmark, and the Great Belt rail tunnel linking Scandinavia with continental Europe. In Australasia: Sydney's Olympic stadiums, Japan's Joetsu Shinkansen high-speed rail line, India's Sardar Sarovar dams, the Surat-Manor toll way project, Calcutta's metro, and Malaysia's Pergau dam. The list ends here only for reasons of space.

Why Cost Overruns and Benefit Shortfalls Are a Problem
Cost overruns and benefit shortfalls of the frequency and size described above are a problem for the following reasons:

- They lead to a Pareto-inefficient allocation of resources, i.e., waste.
- They lead to delays and further cost overruns and benefit shortfalls.
- They destabilize policy, planning, implementation, and operations of projects.
- The problem is getting bigger, because projects get bigger.

Let's consider each point in turn. First, an argument often heard in the planning of large infrastructure projects is that cost and benefit forecasts at the planning stage may be wrong, but if one assumes that forecasts are wrong by the same margin across projects, cost-benefit analysis would still identify the best projects for implementation. The ranking of projects would not be affected by the forecasting errors, according to this argument. However, the large standard deviations shown in tables 1 and 2 falsify the argument. The standard deviations show that cost and benefit estimates are not wrong by the same margin across projects; errors vary extensively and this will affect the ranking of projects. Thus we see that misinformation about costs and benefits at the planning stage is likely to lead to Pareto-inefficiency, because in terms of standard cost-benefit analysis decision makers are likely to implement inferior projects.

Second, cost overruns of the size described above typically lead to delays, because securing additional funding to cover overruns often takes time. In addition, projects may need to be re-negotiated or re-approved when overruns are large, as the data show they often are (Flyvbjerg, 2005a). In a separate study, it was demonstrated that delays in transportation infrastructure implementation are very costly, increasing the percentage construction cost overrun measured in constant prices by 4.64 percentage
points per year of delay incurred after the time of decision to build (Flyvbjerg, Holm, and Buhl, 2004). For a project of, say, US$8 billion—that is the size range of the Channel tunnel and about half the size of Boston's Big Dig—the expected average cost of delay would be approximately $370 million per year, or about $1 million per day. Benefit shortfalls are an additional consequence of delays, because delays result in later opening dates and thus extra months or years without revenues. Because many large infrastructure projects are loan-financed and have long construction periods, they are particularly sensitive to delays, as delays result in increased debt, increased interest payments, and longer payback periods.

Third, large cost overruns and benefit shortfalls tend to destabilize policy, planning, implementation, and operations. For example, after several overruns in the initial phase of the Sydney Opera House, the Parliament of New South Wales decided that every further 10 percent increase in the budget would need their approval. After this decision, the Opera House became a political football needing constant re-approval. Every overrun set off an increasingly menacing debate about the project, in Parliament and outside, with total cost overruns ending at 1,400 percent. The unrest drove the architect off the project, destroyed his career and oeuvre, and produced an Opera House unsuited for opera. Many other projects have experienced similar, if less spectacular, unrest, including the Channel Tunnel, Boston's Big Dig, and Copenhagen's metro.

Finally, as projects grow bigger, the problems with cost overruns and benefit shortfalls also grow bigger and more consequential (Flyvbjerg, Holm, and Buhl, 2004: 12). Some megaprojects are becoming so large in relation to national economies that cost overruns and benefit shortfalls from even a single project may destabilize the finances of a whole country or region. This occurred when the billion-dollar cost overrun on the 2004 Athens Olympics affected the credit rating of Greece and when benefit shortfalls hit Hong Kong’s new $20 billion Chek Lap Kok airport after it opened in 1998. The desire to avoid national fiscal distress has recently become an important driver in attempts at reforming megaproject policy and planning, as we will see later.

Policy Implications

The policy implications of the results presented above are clear:
• Lawmakers, investors, and the public cannot trust information about costs, benefits, and risks of large infrastructure projects produced by promoters and planners of such projects.
• The current way of planning large infrastructure projects is ineffective in conventional economic terms, i.e., it leads to Pareto-inefficient investments.
• There is a strong need for reform in policy and planning for large infrastructure projects.

Before depicting what reform may look like in this expensive and consequential policy area, we will examine the causes of cost overruns and benefit shortfalls.

**Causes of Inaccurate Costs and Benefits**

Three main types of explanation exist that claim to account for inaccuracy in forecasts of costs and benefits: technical, psychological, and political-economic explanations.

*Three Explanations of Cost Overruns and Benefit Shortfalls*

Technical explanations account for cost overruns and benefit shortfalls in terms of unreliable or outdated data, the use of inappropriate forecasting models, honest mistakes, lack of experience on the part of forecasters, etc. This is the most common type of explanation of inaccuracy in forecasts (Ascher, 1978; Flyvbjerg, Holm, and Buhl, 2002, 2005; Morris and Hough, 1987; Vanston and Vanston, 2004; Wachs, 1990). Technical error may be reduced or eliminated by developing better forecasting models, better data, and more experienced forecasters, according to this explanation.

Psychological explanations account for cost overruns and benefit shortfalls in terms of what psychologists call the planning fallacy and optimism bias. Such explanations have been developed by Kahneman and Tversky (1979), Kahneman and Lovallo (1993), and Lovallo and Kahneman (2003). In the grip of the planning fallacy, planners and project promoters make decisions based on delusional optimism rather than on a rational weighting of gains, losses, and probabilities. They overestimate benefits and underestimate costs. They involuntarily spin scenarios of success and overlook the potential for mistakes and miscalculations. As a result, planners and promoters pursue initiatives that are unlikely to come in on budget or on time, or to ever deliver the
expected returns. Overoptimism can be traced to cognitive biases, that is, errors in the way the mind processes information. These biases are thought to be ubiquitous, but their effects can be tempered by simple reality checks, thus reducing the odds that people and organizations will rush blindly into unprofitable investments of money and time.

Political-economic explanations see planners and promoters as deliberately and strategically overestimating benefits and underestimating costs when forecasting the outcomes of projects. They do this in order to increase the likelihood that it is their projects, and not the competition's, that gain approval and funding. Political-economic explanations have been set forth by Flyvbjerg, Holm, and Buhl (2002, 2005) and Wachs (1989, 1990). According to such explanations planners and promoters purposely spin scenarios of success and gloss over the potential for failure. Again, this results in the pursuit of ventures that are unlikely to come in on budget or on time, or to deliver the promised benefits. Strategic misrepresentation can be traced to political and organizational pressures, for instance competition for scarce funds or jockeying for position, and it is rational in this sense. If we now define a lie in the conventional fashion as making a statement intended to deceive others (Bok, 1979: 14; Cliffe et al., 2000: 3), we see that deliberate misrepresentation of costs and benefits is lying, and we arrive at one of the most basic explanations of lying that exists: Lying pays off, or at least political and economic agents believe it does. Where there is political pressure there is misrepresentation and lying, according to this explanation, but misrepresentation and lying can be moderated by measures of accountability.

Validity of Explanations
How well does each of the three explanations of forecasting inaccuracy—technical, psychological, and political-economic—account for the data on cost overruns and benefit shortfalls presented earlier? This is the question to be answered in this section.

Technical explanations have, as mentioned, gained widespread credence among forecasters and planners (Ascher, 1978; Flyvbjerg, Holm, and Buhl, 2002, 2005; Vanston and Vanston, 2004). It turns out, however, that such credence could mainly be upheld because until now samples have been too small to allow tests by statistical methods. The data presented above, which come from the first large-sample study in the field, lead us to reject technical explanations of forecasting inaccuracy. Such explanations do not fit the data well. First, if misleading forecasts were truly caused by technical inadequacies, simple mistakes, and inherent problems with predicting the future, we would expect
a less biased distribution of errors in forecasts around zero. In fact, we have found with high statistical significance that for four out of five distributions of forecasting errors, the distributions have a mean statistically different from zero. Only the data for inaccuracy in road traffic forecasts have a statistical distribution that seem to fit with explanations in terms of technical forecasting error. This seeming fit is currently being tested in a separate study (Næss, Flyvbjerg, and Buhl, 2006). Second, if imperfect techniques, inadequate data, and lack of experience were main explanations of inaccuracies, we would expect an improvement in accuracy over time, since in a professional setting errors and their sources would be recognized and addressed through the refinement of data collection, forecasting methods, etc. Substantial resources have in fact been spent over several decades on improving data and methods. Still the data show that this has had no effect on the accuracy of forecasts. Technical factors, therefore, do not appear to explain the data. It is not so-called forecasting “errors” or their causes that need explaining. It is the fact that in a large majority of cases, costs are underestimated and benefits overestimated. One may agree with proponents of technical explanations that it is, for example, impossible to predict for the individual project exactly which geological, environmental, or safety problems will appear and make costs soar. But one must maintain that it is possible to predict the risk, based on experience from other projects, that some such problems will haunt a project and how this will affect costs. One must also maintain that such risk can and should be accounted for in forecasts of costs, but typically is not. For technical explanations to be valid, they would have to explain why forecasts are so consistent in ignoring cost and benefit risks over time, location, and project type.

Psychological explanations better fit the data. The existence of optimism bias in planners and promoters would result in actual costs being higher and actual benefits being lower than those forecasted. Consequently, the existence of optimism bias would be able to account, in whole or in part, for the peculiar bias found in most of the data. Interestingly, however, when you ask forecasters about causes for forecasting inaccuracies in actual forecasts, they do not mention optimism bias as a main cause of inaccuracy (Flyvbjerg, Holm, and Buhl, 2005: 138-140). This could of course be because optimism bias is unconscious and thus not reflected by forecasters. After all, there is a large body of experimental evidence for the existence of optimism bias (Buehler et al., 1994; Buehler, Griffin, and MacDonald, 1997; Newby-Clark et al. 2002). However, the experimental data are mainly from simple, non-professional settings. This is a problem for psychological explanations, because it leaves open the question of whether such
explanations are general and apply beyond the simple settings studied. Optimism bias would be an important and credible explanation of underestimated costs and overestimated benefits in infrastructure forecasting if estimates were produced by inexperienced forecasters, i.e., persons who were estimating costs and benefits for the first or second time and who were thus unknowing about the realities of infrastructure building and were not drawing on the knowledge and skills of more experienced colleagues. Such situations may exist and may explain individual cases of inaccuracy. But given the fact that in modern society it is a defining characteristic of professional expertise that it is constantly tested--through scientific analysis, critical assessment, and peer review--in order to root out bias and error, it seems unlikely that a whole profession of forecasting experts would continue to make the same mistakes decade after decade instead of learning from their actions. Learning would result in the reduction, if not elimination, of optimism bias, which would then result in estimates becoming more accurate over time. But the data clearly show that this has not happened. The profession of forecasters would indeed have to be an optimistic--and non-professional--group to keep their optimism bias throughout the 70-year period the study covers for costs, and the 30-year period covered for demand, and not learn that they were deceiving themselves and others by underestimating costs and overestimating benefits. This would account for the data, but is not a credible explanation. Therefore, on the basis of the data, we are led to reject optimism bias as a primary cause of cost underestimation and benefit overestimation.

Political-economic explanations and strategic misrepresentation account well for the systematic underestimation of costs and overestimation of benefits found in the data. A strategic estimate of costs would be low, resulting in cost overrun, whereas a strategic estimate of benefits would be high, resulting in benefit shortfalls. A key question for explanations in terms of strategic misrepresentation is whether estimates of costs and benefits are intentionally biased to serve the interests of promoters in getting projects started. This question raises the difficult issue of lying. Questions of lying are notoriously hard to answer, because a lie is making a statement intended to deceive others, and in order to establish whether lying has taken place, one must therefore know the intentions of actors. For legal, economic, moral, and other reasons, if promoters and planners have intentionally manipulated estimates of costs and benefits to get a project started, they are unlikely to formally tell researchers or others that this is the case. Despite such problems, two studies exist that succeeded in getting forecasters to talk about strategic misrepresentation (Flyvbjerg and Cowi, 2004; Wachs 1990).
Flyvbjerg and Cowi (2004) interviewed public officials, planners, and consultants who had been involved in the development of large UK transportation infrastructure projects. A planner with a local transportation authority is typical of how respondents explained the basic mechanism of cost underestimation:

“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.”

Experienced professionals like the interviewee know that outturn costs will be higher than estimated costs, but because of political pressure to secure funding for projects they hold back this knowledge.

Similarly, an interviewee explained the basic mechanism of benefit overestimation:

“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, also because the project promoters know that their project is up against other projects and competing for scarce resources.”

Such a focus on benefits and disregard of risks and robustness may consist, for instance, in the discounting of spatial assimilation problems described by Priemus (forthcoming). Competition between projects and authorities creates political and organizational pressures that in turn create an incentive structure that makes it rational for project promoters to emphasize benefits and de-emphasize costs and risks. A project that looks highly beneficial on paper is more likely to get funded than one that does not.

Specialized private consultancy companies are typically engaged to help develop project proposals. In general, the interviewees found that consultants showed high professional standard and integrity. But interviewees also found that consultants appeared to focus on justifying projects rather than critically scrutinizing them. A project manager explained:
"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."

The consultants interviewed confirmed that appraisals often focused more on benefits than on costs. But they said this was at the request of clients and that for specific projects discussed "there was an incredible rush to see projects realized."

One typical interviewee saw project approval as "passing the test" and precisely summed up the rules of the game like this:

"It’s all about passing the test [of project approval]. You are in, when you are in. It means that there is so much focus on showing the project at its best at this stage."

In sum, the UK study shows that strong interests and strong incentives exist at the project approval stage to present projects as favorably as possible, that is, with benefits emphasized and costs and risks deemphasized. Local authorities, local developers and land owners, local labor unions, local politicians, local officials, local MPs, and consultants all stand to benefit from a project that looks favorable on paper and they have little incentive to actively avoid bias in estimates of benefits, costs, and risks. National bodies, like certain parts of the Department for Transport and the Ministry of Finance who fund and oversee projects, may have an interest in more realistic appraisals, but so far they have had little success in achieving such realism, although the situation may be changing with the initiatives to curb bias set out in HM Treasury (2003).

The second study was carried out by Martin Wachs (1990, 1986). Wachs interviewed public officials, consultants, and planners who had been involved in transit planning cases in the US. He found that a pattern of highly misleading forecasts of costs and patronage could not be explained by technical errors, honest mistakes, or inadequate methods. In case after case, planners, engineers, and economists told Wachs that they had had to "revise" their forecasts many times because they failed to satisfy their superiors. The forecasts had to be cooked in order to produce numbers that were dramatic enough to gain federal support for the projects whether or not they could be fully justified on technical grounds. Wachs (1990: 144) recounts from his interviews:
"One young planner, tearfully explained to me that an elected county supervisor had asked her to estimate the patronage of a possible extension of a light-rail (streetcar) line to the downtown Amtrak station. When she carefully estimated that the route might carry two to three thousand passengers per day, the supervisor directed her to redo her calculations in order to show that the route would carry twelve to fifteen thousand riders per day because he thought that number necessary to justify a federal grant for system construction. When she refused, he asked her superior to remove her from the project, and to get someone else to 'revise' her estimates."

In another typical case of cost underestimation and benefit overestimation, Wachs (1990: 144-145) gives the following account:

"a planner admitted to me that he had reluctantly but repeatedly adjusted the patronage figures upward, and the cost figures downward to satisfy a local elected official who wanted to compete successfully for a federal grant. Ironically, and to the chagrin of that planner, when the project was later built, and the patronage proved lower and the costs higher than the published estimates, the same local politician was asked by the press to explain the outcome. The official's response was to say, 'It's not my fault; I had to rely on the forecasts made by our staff, and they seem to have made a big mistake here'."

Like in the UK study above, Wachs specifically spoke with consultants. He found, as one consultant put it, that "success in the consulting business requires the forecaster to adjust results to conform with the wishes of the client," and clients typically wish to see costs underestimated and benefits overestimated (1990: 151-152).

On the basis of his pioneering study, Wachs (1990: 145) concludes that forecasts of costs and benefits are presented to the public as instruments for deciding whether or not a project is to be undertaken, but they are actually instruments for getting public funds committed to a favored project. Wachs (1990: 146, 1986: 28) talks of "nearly universal abuse" of forecasting in this context, and he finds no indication that it takes place only in transit planning; it is common in all sectors of the economy where forecasting routinely plays an important role in policy debates, according to Wachs.

In conclusion, the UK and US studies arrive at results that are basically similar.
Both studies account well for existing data on cost underestimation and benefit overestimation. Both studies falsify the notion that in situations with high political and organizational pressure the lowballing of costs and highballing of benefits is caused by non-intentional technical error or optimism bias. Both studies support the view that in such situations promoters and forecasters intentionally use the following formula in order to secure approval and funding for their projects:

\[
\text{Underestimated costs} + \text{Overestimated benefits} = \text{Project approval}
\]

Using this formula, and thus "showing the project at its best" as one interviewee said above, results in an inverted Darwinism, i.e., the "survival of the unfittest." It is not the best projects that get implemented, but the projects that look best on paper. And the projects that look best on paper are the projects with the largest cost underestimates and benefit overestimates, other things being equal. But these are the worst, or "unfittest," projects in the sense that they are the very projects that will encounter most problems during construction and operations in terms of the largest cost overruns, benefit shortfalls, and risks of non-viability. The projects have been designed like that.

Towards Falsifiable Theories
Figure 1 schematically sums up the relationship between the two types of explanation that best fit the data, psychological and political-economic explanations. Psychological explanations, which account for the data in terms of optimism bias, have their relative merit in situations where political and organizational pressures are absent or low, whereas such explanations hold less power in situations where political pressures are high. Conversely, political-economic explanations, which account for the data in terms of strategic misrepresentation, have their relative merit where political and organizational pressures are high--this being common for large, high-profile, public projects with many and powerful stakeholders--while these explanations become immaterial when such pressures are not present.

[Figure 1 app. here]
Both types of explanation may be falsified or verified by specific, empirical evidence. Karl Popper used the famous statement "all swans are white" to illustrate what a falsifiable proposition is, and Popper argued that the observation of a black swan would falsify this statement. In the same manner the statement "all forecasts are biased due to optimism bias"--which was the common view until recently (Kahneman and Lovallo, 2003; Flyvbjerg, 2003; HM Treasury, 2003)--may be falsified by the observation of a forecast that is biased due to other reasons, for instance strategic misrepresentation. An example of such falsification is the evidence uncovered by the interviews presented above with public officials, planners, and consultants who have worked on large transportation infrastructure projects. This evidence falsifies the notion that optimism bias is a sole and adequate explanation for biased forecasts of costs and benefits in these projects. However, rather than compete, the two types of explanation complement each other: one is strong where the other is weak, and both explanations are necessary to understand the phenomenon at hand--the pervasiveness of inaccuracy and risk in decision making--and how to curb it.

The two types of explanation also point directly to two different types of solutions to the problem of bias in forecasting. First, in situations where it has been established that optimism bias is the main cause of inaccuracy we may assume that managers and forecasters are making honest mistakes and have an interest in improving accuracy. Here improved methods may help. Forecasters will welcome such methods, because no one in this situation has reason to be against a methodology that will improve their forecasts.

Second, when strategic misrepresentation is the main cause of inaccuracy, differences between estimated and actual costs and benefits are best explained by political and organizational pressures. Here forecasters would still need better methods if accuracy were to be improved, but forecasters may not be interested in this because inaccuracy is deliberate. Biased forecasts serve strategic purposes that dominate the commitment to accuracy and truth.

Consider, for example, city managers with responsibility for estimating costs and benefits of urban rail projects. Here, the assumption of innocence regarding outcomes typically cannot be upheld. Cities compete intensely for approval and for scarce national funds for such projects, and pressures are strong to present projects as favorably as possible, that is, with low costs and high benefits, in order to beat the competition. There is no incentive for the individual city to debias its forecasts, but quite the opposite. Unless all other cities also debias, the individual city would lose out in the competition.
for funds. This is the type of situation documented in Flyvbjerg and Cowi (2004: 36-58). The result is the same as in the case of optimism: actors promote ventures that are unlikely to perform as promised. But the causes are different as are possible cures.

The existence of strategic misrepresentation does not exclude the simultaneous existence of optimism bias, and vice versa. In fact, it is realistic to expect such co-existence in forecasting in large and complex projects and organizations. This underscores that improved forecasting methods must go hand in hand with measures of accountability if attempts to arrive at more accurate forecasts is to be effective. Such attempts are the concern of the following section.

**Cures: The Need for Reform in Megaproject Policy and Planning**

As should be clear, the planning and implementation of large infrastructure projects stand in need of reform. Less deception and more honesty are needed in the estimation of costs and benefits if better projects are to be implemented. This is not to say that costs and benefits are or should be the only basis for deciding whether to build large infrastructure projects. Clearly, forms of rationality other than economic rationality are at work in most projects and are balanced in the broader frame of public decision making. But the costs and benefits of large infrastructure projects often run in the hundreds of millions of dollars, with risks correspondingly high. Without knowledge of such risks, decisions are likely to be flawed.

When contemplating what planners can do to help reform come about, we need to distinguish between two fundamentally different situations: (1) planners and promoters consider it important to get forecasts of costs, benefits, and risks right, and (2) planners and promoters do not consider it important to get forecasts right, because optimistic forecasts are seen as a necessary means to getting projects started. The first situation is the easier one to deal with and here better methodology will go a long way in improving planning and decision making. The second situation is more difficult, and more common as we saw above. Here changed incentives are essential in order to reward honesty and punish deception, where today's incentives often do the opposite.

Thus two main measures of reform may be identified: (1) better forecasting methods, and (2) improved incentive structures, with the latter being more important.
Better Methods: Reference Class Forecasting

If planners genuinely consider it important to get forecasts right, it is recommended they use a new forecasting method called "reference class forecasting" to reduce inaccuracy and bias. This method was originally developed to compensate for the type of cognitive bias in human forecasting that Princeton psychologist Daniel Kahneman found in his Nobel prize-winning work on bias in economic forecasting (Kahneman, 1994; Kahneman and Tversky, 1979). Reference class forecasting has proven more accurate than conventional forecasting. Flyvbjerg and Cowi (2004) developed the first instance of practical reference class forecasting. Later, the UK Department for Transport and HM Treasury decided to employ the method in the UK as part of project appraisal. They placed guidelines on how to use the method on their websites, for local government and consultants to follow. In April 2005, based on a study by Flyvbjerg, Holm, and Buhl (2005), the American Planning Association (2005) officially endorsed reference class forecasting:

"APA encourages planners to use reference class forecasting in addition to traditional methods as a way to improve accuracy. The reference class forecasting method is beneficial for non-routine projects such as stadiums, museums, exhibit centers, and other local one-off projects. Planners should never rely solely on civil engineering technology as a way to generate project forecasts."

For reasons of space, here only an outline of the method is presented, based mainly on Lovallo and Kahneman (2003) and Flyvbjerg (2003).

Reference class forecasting consists in taking a so-called "outside view" on the particular project being forecast. The outside view is established on the basis of information from a class of similar projects. The outside view does not try to forecast the specific uncertain events that will affect the particular project, but instead places the project in a statistical distribution of outcomes from this class of reference projects. Reference class forecasting requires the following three steps for the individual project:

1. Identification of a relevant reference class of past projects. The class must be broad enough to be statistically meaningful but narrow enough to be truly comparable with the specific project.
(2) Establishing a probability distribution for the selected reference class. This requires access to credible, empirical data for a sufficient number of projects within the reference class to make statistically meaningful conclusions.

(3) Comparing the specific project with the reference class distribution, in order to establish the most likely outcome for the specific project.

Daniel Kahneman relates the following story about curriculum planning to illustrate reference class forecasting in practice (Lovallo and Kahneman 2003: 61). This example is used, because similar examples do not exist as yet in the field of infrastructure planning. Some years ago, Kahneman was involved in a project to develop a curriculum for a new subject area for high schools in Israel. The project was carried out by a team of academics and teachers. In time, the team began to discuss how long the project would take to complete. Everyone on the team was asked to write on a slip of paper the number of months needed to finish and report the project. The estimates ranged from 18 to 30 months. One of the team members—a distinguished expert in curriculum development—was then posed a challenge by another team member to recall as many projects similar to theirs as possible and to think of these projects as they were in a stage comparable to their project. "How long did it take them at that point to reach completion?", the expert was asked. After a while he answered that not all the comparable teams he could think of ever did complete their task. About 40 percent of them eventually gave up. Of those remaining, the expert could not think of any that completed their task in less than seven years, nor of any that took more than ten. The expert was then asked if he had reason to believe that the present team was more skilled in curriculum development than the earlier ones had been. The expert said no, he did not see any relevant factor that distinguished this team favorably from the teams he had been thinking about. His impression was that the present team was slightly below average in terms of resources and potential. The wise decision at this point would probably have been for the team to break up, according to Kahneman. Instead, the members ignored the pessimistic information and proceeded with the project. They finally completed the project eight years later, and their efforts went largely wasted—the resulting curriculum was rarely used.

In this example, the curriculum expert made two forecasts for the same problem and arrived at very different answers. The first forecast was the inside view; the second was the outside view, or the reference class forecast. The inside view is the one that the expert and the other team members adopted. They made forecasts by focusing tightly on the case at hand, considering its objective, the resources they brought to it, and the
obstacles to its completion. They constructed in their minds scenarios of their coming progress and extrapolated current trends into the future. The resulting forecasts, even the most conservative ones, were overly optimistic. The outside view is the one provoked by the question to the curriculum expert. It completely ignored the details of the project at hand, and it involved no attempt at forecasting the events that would influence the project's future course. Instead, it examined the experiences of a class of similar projects, laid out a rough distribution of outcomes for this reference class, and then positioned the current project in that distribution. The resulting forecast, as it turned out, was much more accurate.

Similarly—to take an example from the work with developing reference class forecasting for practical infrastructure planning—planners in a city preparing to build a new subway would, first, establish a reference class of comparable projects. This could be the relevant rail projects from the sample used for this article. Through analyses the planners would establish that the projects included in the reference class were indeed comparable. Second, if the planners were concerned, for example, with getting construction cost estimates right, they would then establish the distribution of outcomes for the reference class regarding the accuracy of construction cost forecasts. Figure 2 shows what this distribution looks like for a reference class relevant to building subways in the UK, developed by Flyvbjerg and Cowi (2004: 23) for the UK Department for Transport.

Third, the planners would compare their subway project to the reference class distribution. This would make it clear to the planners that unless they have reason to believe they are substantially better forecasters and planners than their colleagues who did the forecasts and planning for projects in the reference class, they are likely to grossly underestimate construction costs. Finally, planners would then use this knowledge to adjust their forecasts for more realism. Figure 3 shows what such adjustments are for the UK situation. More specifically, Figure 3 shows that for a forecast of construction costs for a rail project, which has been planned in the manner that such projects are usually planned, i.e., like the projects in the reference class, this forecast would have to be adjusted upwards by 40 percent, if investors were willing to accept a risk of cost overrun of 50 percent. If investors were willing to accept a risk of overrun of only 10 percent, the uplift would have to be 68 percent. For a rail project initially estimated at, say £4 billion, the uplifts for the 50 and 10 percent levels of risk of cost overrun would be £1.6 billion and £2.7 billion, respectively.

[Figures 2 and 3 app. here]
The contrast between inside and outside views has been confirmed by systematic research (Gilovich, Griffin, and Kahneman, 2002). The research shows that when people are asked simple questions requiring them to take an outside view, their forecasts become significantly more accurate. However, most individuals and organizations are inclined to adopt the inside view in planning major initiatives. This is the conventional and intuitive approach. The traditional way to think about a complex project is to focus on the project itself and its details, to bring to bear what one knows about it, paying special attention to its unique or unusual features, trying to predict the events that will influence its future. The thought of going out and gathering simple statistics about related cases seldom enters a planner's mind. This is the case in general, according to Lovallo and Kahneman (2003: 61-62). And it is certainly the case for cost and benefit forecasting in large infrastructure projects. Despite the many forecasts reviewed for the present study, not a single genuine reference class forecast of costs and benefits has been identified.\(^7\)

While understandable, planners' preference for the inside view over the outside view is unfortunate. When both forecasting methods are applied with equal skill, the outside view is much more likely to produce a realistic estimate. That is because it bypasses cognitive and political biases such as optimism bias and strategic misrepresentation and cuts directly to outcomes. In the outside view planners and forecasters are not required to make scenarios, imagine events, or gauge their own and others' levels of ability and control, so they cannot get all these things wrong. Human bias is bypassed. Undoubtedly the outside view, being based on historical precedent, may fail to predict extreme outcomes, that is, those that lie outside all historical precedents. But for most projects, the outside view will produce more accurate results. In contrast, a focus on inside details is the road to inaccuracy.

The comparative advantage of the outside view is most pronounced for non-routine projects, understood as projects that planners and decision makers in a certain locale have never attempted before—like building an urban rail system in a city for the first time, or a new major bridge or tunnel where none existed before. It is in the planning of such new efforts that the biases toward optimism and strategic misrepresentation are likely to be largest. To be sure, choosing the right reference class of comparative past projects becomes more difficult when planners are forecasting initiatives for which precedents are not easily found, for instance the introduction of new and unfamiliar technologies. However, most large infrastructure projects are both non-routine locally and use well-known technologies. Such projects are, therefore, particularly likely to
benefit from the outside view and reference class forecasting. The same holds for concert halls, museums, stadiums, exhibition centers, and other local one-off projects.

**Improved Incentives: Public and Private Sector Accountability**

In the present section we consider the situation where planners and other influential actors do not find it important to get forecasts right and where planners, therefore, do not help to clarify and mitigate risk but, instead, generate and exacerbate it. Here planners are part of the problem, not the solution. This situation may need some explication, because it possibly sounds to many like an unlikely state of affairs. After all, it may be agreed that planners ought to be interested in being accurate and unbiased in forecasting. It is even stated as an explicit requirement in the American Institute of Certified Planners' Code of Ethics and Professional Conduct that "A planner must strive to provide full, clear and accurate information on planning issues to citizens and governmental decision-makers" (American Planning Association, 1991: A.3). The British Royal Town Planning Institute has laid down similar obligations for its members (Royal Town Planning Institute, 2001).

However, the literature is replete with things planners and planning "must" strive to do, but which they don't. Planning must be open and communicative, but often it is closed. Planning must be participatory and democratic, but often it is an instrument of domination and control. Planning must be about rationality, but often it is about power (Flyvbjerg, 1998; Watson, 2003). This is the "dark side" of planning and planners identified by Flyvbjerg (1996) and Yiftachel (1998), which is remarkably underexplored by planning researchers and theorists.

Forecasting, too, has its dark side. It is here that "planners lie with numbers," as Wachs (1989) aptly put it. Planners on the dark side are busy not with getting forecasts right and following the AICP Code of Ethics but with getting projects funded and built. And accurate forecasts are often not an effective means for achieving this objective. Indeed, accurate forecasts may be counterproductive, whereas biased forecasts may be effective in competing for funds and securing the go-ahead for construction. "The most effective planner," says Wachs (1989: 477), "is sometimes the one who can cloak advocacy in the guise of scientific or technical rationality." Such advocacy would stand in direct opposition to AICP's ruling that "the planner's primary obligation [is] to the public interest" (American Planning Association, 1991: B.2). Nevertheless, seemingly rational forecasts that underestimate costs and overestimate benefits have long been an
established formula for project approval as we saw above. Forecasting is here mainly another kind of rent-seeking behavior, resulting in a make-believe world of misrepresentation which makes it extremely difficult to decide which projects deserve undertaking and which do not. The consequence is, as even one of the industry's own organs, the Oxford-based Major Projects Association, acknowledges, that too many projects proceed that should not. One may add that many projects don't proceed that probably should, had they not lost out to projects with "better" misrepresentation (Flyvbjerg, Holm, and Buhl, 2002).

In this situation, the question is not so much what planners can do to reduce inaccuracy and risk in forecasting, but what others can do to impose on planners the checks and balances that would give planners the incentive to stop producing biased forecasts and begin to work according to their Code of Ethics. The challenge is to change the power relations that govern forecasting and project development. Better forecasting techniques and appeals to ethics won't do here; institutional change with a focus on transparency and accountability is necessary.

As argued in Flyvbjerg, Bruzelius, and Rothengatter (2003), two basic types of accountability define liberal democracies: (1) public sector accountability through transparency and public control, and (2) private sector accountability via competition and market control. Both types of accountability may be effective tools to curb planners' misrepresentation in forecasting and to promote a culture which acknowledges and deals effectively with risk. In order to achieve accountability through transparency and public control, the following would be required as practices embedded in the relevant institutions (the full argument for the measures may be found in Flyvbjerg, Bruzelius, and Rothengatter, 2003, chapters 9-11):

- National-level government should not offer discretionary grants to local infrastructure agencies for the sole purpose of building a specific type of infrastructure. Such grants create perverse incentives. Instead, national government should simply offer "infrastructure grants" or "transportation grants" to local governments, and let local political officials spend the funds however they choose to, but make sure that every dollar they spend on one type of infrastructure reduces their ability to fund another.

- Forecasts should be made subject to independent peer review. Where large amounts of taxpayers' money are at stake, such review may be 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, who have the independence and expertise to produce such reviews. Other types of independent review bodies may be established, for instance within national departments of finance or with relevant professional bodies.

- Forecasts should be benchmarked against comparable forecasts, for instance using reference class forecasting as described in the previous section.
- Forecasts, peer reviews, and benchmarkings should be made available to the public as they are produced, including all relevant documentation.
- Public hearings, citizen juries, and the like should be organized to allow stakeholders and civil society to voice criticism and support of forecasts. Knowledge generated in this way should be integrated in planning and decision making.
- Scientific and professional conferences should be organized where forecasters would present and defend their forecasts in the face of colleagues’ scrutiny and criticism.
- Projects with inflated benefit-cost ratios should be reconsidered and stopped if recalculated costs and benefits do not warrant implementation. Projects with realistic estimates of benefits and costs should be rewarded.
- Professional and occasionally even criminal penalties should be enforced for planners and forecasters who consistently and foreseeably produce deceptive forecasts. An example of a professional penalty would be the exclusion from one’s professional organization if one violates its code of ethics. An example of a criminal penalty would be punishment as the result of prosecution before a court or similar legal set-up, for instance where deceptive forecasts have led to substantial mismanagement of public funds (Garett and Wachs, 1996). Malpractice in planning should be taken as seriously as it is in other professions. Failing to do this amounts to not taking the profession of planning seriously.

In order to achieve accountability in forecasting via *competition and market control*, the following would be required, again as practices that are both embedded in and enforced by the relevant institutions:

- The decision to go ahead with a project should, where at all possible, 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. This should be
required whether projects pass the market test or not, that is, whether projects are subsidized or not or provided for social justice reasons or not. Private lenders, shareholders, and stock market analysts would produce their own forecasts or would critically monitor existing ones. If they were wrong about the forecasts, they and their organizations would be hurt. The result would be more realistic forecasts and reduced risk.

- Full public financing or full financing with a sovereign guarantee should be avoided.
- Forecasters and their organizations should share financial responsibility for covering cost overruns and benefit shortfalls resulting from misrepresentation and bias in forecasting.
- The participation of risk capital should not mean that government gives up or reduces control of the project. On the contrary, it means that government can more effectively play the role it should be playing, namely keeping the project at arms length as the ordinary citizen's guarantor for ensuring concerns about safety, environment, risk, and a proper use of public funds.

Whether projects are public, private, or public-private, they should be vested in one and only one project organization with a strong governance framework. The project organization may be a company or not, public or private, or a mixture. What is important is that this organization enforces accountability vis-à-vis contractors, operators, etc., and that, in turn, the directors of the organization are held accountable for any cost overruns, benefits shortfall, faulty designs, unmitigated risks, etc. that may occur during project planning, implementation, and operations.

If the institutions with responsibility for developing and building major infrastructure projects would effectively implement, embed, and enforce such measures of accountability, then the misrepresentation in cost, benefit, and risk estimates, which is widespread today, may be mitigated. If this is not done, misrepresentation is likely to continue, and the allocation of funds for infrastructure is likely to continue to be wasteful and undemocratic.
Validity of Results Across Types of Infrastructure

The cost data included in the study of transportation infrastructure projects presented above was compared with cost data for other project types in other studies including power plants, dams, public buildings, sports facilities, water distribution, oil and gas extraction, information technology systems, aerospace systems, and weapons systems (Arditi et al., 1985; Blake et al., 1976; Byggeriets Evaluering Center, 2006; Canaday, 1980; Delaney and Eckstein, 2003; Department of Energy Study Group, 1975; Dlakwa and Culpin, 1990; Fraser, 1990; Hall, 1980; Healey, 1964; Henderson, 1977; Hufschmidt and Gerin, 1970; Merewitz, 1973b; Merrow, 1988; Morris and Hough, 1987; Mott MacDonald, 2002; National Audit Office, 2004; National Audit Office of Denmark, 2000; National Audit Office of Norway, 2001; Norwegian Public Investigatory Commissions, 1999; Preuss, 2000; Rich, 2000; Standish Group, 2004; World Bank, 1994; World Commission on Dams, 2000). None of these studies are as comprehensive and systematic as the study of transportation infrastructure presented here. Nevertheless, the data from the studies indicate that other types of projects are as prone to cost overrun as are transportation infrastructure projects. The data also indicate that cost overrun for other projects have neither increased nor decreased historically, and that overrun is common in both First- and Third-World countries, as is the case for transportation infrastructure. Therefore, transportation infrastructure appears to be not a unique case of cost overrun but, instead, one example among many of a more general phenomenon, namely cost overrun in infrastructure policy and planning.

Data on benefit shortfalls are more scarce than data on cost overrun. However, some of the studies mentioned in the previous paragraph contain data on benefits (Delaney and Eckstein, 2003; Hall, 1980; Merrow, 1988; Mott MacDonald, 2002; Preuss, 2000; Rich, 2000; Standish Group, 2004; World Bank, 1994; World Commission on Dams, 2000). With the reservations that apply to sparse data, again the available data indicate that large and frequent benefit shortfalls are not unique to transportation infrastructure but is a more widespread problem, common to infrastructure development in general.

Towards Better Practice

Fortunately, after decades of widespread mismanagement in megaproject policy and planning, signs of improvement have recently appeared. The conventional consensus that
deception is an acceptable way of getting projects started is under attack, as will be apparent from the examples below. This is in part because democratic governance is generally getting stronger around the world. The Enron scandal and its successors have triggered a war on corporate deception that is spilling over into government with the same objective: to reduce financial waste and promote good governance. Although progress is slow, good governance is gaining a foothold even in megaproject policy and planning. The conventional consensus is also under attack for the practical reason mentioned earlier that the largest projects are now so big in relation to national economies that cost overruns, benefit shortfalls, and risks from even a single project may destabilize the finances of a whole country or region, as happened in Greece and Hong Kong. Lawmakers and governments begin to see that national fiscal distress is too high a price to pay for the conventional way of planning and designing large projects. The main drive for reform comes from outside the agencies and industries conventionally involved in infrastructure development, which increases the likelihood of success.

In 2003 the Treasury of the United Kingdom required, for the first time, that all ministries develop and implement procedures for large public projects that will curb what the Treasury calls--with true British civility--"optimism bias." Funding will be unavailable for projects that do not take into account this bias, and methods have been developed for how to do this (Mott MacDonald, 2002; HM Treasury, 2003; Flyvbjerg and Cowi, 2004). In the Netherlands in 2004, the Parliamentary Committee on Infrastructure Projects for the first time conducted extensive public hearings to identify measures that will limit the misinformation about large infrastructure projects given to the Parliament, public, and media (Tijdelijke Commissie Infrastructuurprojecten, 2004). In Boston, the government sued to recoup funds from contractor overcharges for the Big Dig related to cost overruns. More governments and parliaments are likely to follow the lead of the UK, the Netherlands, and Boston in coming years. It’s too early to tell whether the measures they implement will ultimately be effective. It seems unlikely, however, that the forces that have triggered the measures will be reversed, and it is those forces that reform-minded groups need to support and develop in order to curb deception and waste. This is the "tension-point" where convention meets reform, power-balances change, and new things may happen.

The key weapons in the war on deception and waste is accountability and critical questioning. The professional expertise of planners, engineers, architects, economists, and administrators is certainly indispensable to constructing the infrastructures that make society work. Studies show, however, that the claims about costs, benefits, and risks
made by these groups usually cannot be trusted and should be carefully examined by independent specialists and organizations. The same holds for claims made by project-promoting politicians and officials. Institutional checks and balances--including financial, professional, or even criminal penalties for consistent and unjustifiable biases in claims and estimates of costs, benefits, and risks--should be developed and employed. The key principle is that the cost of making a wrong forecast should fall on those making the forecast, a principle often violated today.

Some of the public-private partnerships currently emerging in large infrastructure projects contain more and better checks and balances than previous institutional setups, as has been demonstrated by the National Audit Office (2003). This is a step in the right direction but should be no cause for repose. The task of turning around megaproject policy and planning is so difficult that all available measures for improvement must be employed. The conventional mode of planning and designing infrastructure has long historical roots and is deeply ingrained in professional and institutional practices. It would be naive to think it is easily toppled. Given the stakes involved--saving taxpayers from billions of dollars of waste, protecting citizens’ trust in democracy and the rule of law, avoiding the destruction of spatial and environmental assets--this shouldn’t deter us from trying.
References


Byggeriets Evaluerings Center, 2006, *Budgetoverskridelser på byggeprojekter: Et forstudie for Fonden Realdania* (Copenhagen and Aalborg: Author)


Flyvbjerg, Bent, 2005b, Measuring Inaccuracy in Travel Demand Forecasting:
Methodological Considerations Regarding Ramp Up and Sampling, *Transportation Research A*, vol. 39, no. 6, pp 522-530


Hall, Peter, 1980, Great Planning Disasters (Harmondsworth: Penguin Books)


Lovallo, Dan and Daniel Kahneman, 2003, Delusions of Success: How Optimism


National Audit Office of Denmark, 2000, *Beretning til statsrevisorerne om gennemførelse af statslige edb-projekter*, RB B202/00 (Copenhagen: Author)


Uncomfortable? *Journal of Personality and Social Psychology*, 82, pp 157-166


Priemus, Hugo, forthcoming, Design of Large Infrastructure Projects: Disregarded Alternatives and Issues of Spatial Planning


Table 1: Inaccuracy of transportation project cost estimates by type of project. Constant prices.

<table>
<thead>
<tr>
<th>Type of project</th>
<th>No. of cases (N)</th>
<th>Avg. cost overrun %</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail</td>
<td>58</td>
<td>44.7</td>
<td>38.4</td>
</tr>
<tr>
<td>Bridges and tunnels</td>
<td>33</td>
<td>33.8</td>
<td>62.4</td>
</tr>
<tr>
<td>Road</td>
<td>167</td>
<td>20.4</td>
<td>29.9</td>
</tr>
</tbody>
</table>
Table 2: Inaccuracy in forecasts of rail passenger and road vehicle traffic.

<table>
<thead>
<tr>
<th>Type of project</th>
<th>No. of cases (N)</th>
<th>Avg. inaccuracy %</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rail</td>
<td>25</td>
<td>-51.4</td>
<td>28.1</td>
</tr>
<tr>
<td>Road</td>
<td>183</td>
<td>9.5</td>
<td>44.3</td>
</tr>
</tbody>
</table>
Figure 1: Explanatory power of optimism bias and strategic misrepresentation, respectively, in accounting for forecasting inaccuracy as function of political and organizational pressure.
Figure 2: Inaccuracy of construction cost forecasts for rail projects in reference class. Average cost increase is indicated for non-UK and UK projects, separately. Constant prices.
Figure 3: Required adjustments to cost estimates for UK rail projects as function of the maximum acceptable level of risk for cost overrun. Constant prices.
Notes

1 A “megaproject” is here defined as the most expensive infrastructure and investment projects that are carried out today, typically at costs per project from several hundred million to several billion dollars.

2 All costs are construction costs measured in constant prices. Cost overrun, also sometimes called “cost increase” or “cost escalation,” is measured according to international convention as actual out-turn costs minus estimated costs in percent of estimated costs. Actual 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. For reasons explained in Flyvbjerg, Holm, and Buhl (2002) and Flyvbjerg (2005b) the figures for cost overrun presented here must be considered conservative—ideally financing costs, operating costs, and maintenance costs would also be included in a study of costs. It is difficult, however, to find valid, reliable, and comparable data on these types of costs across a large number of projects.

3 Following international convention, inaccuracy is measured as actual traffic minus estimated traffic in percent of estimated traffic. Rail traffic is measured as number of passengers; road traffic as number of vehicles. The base year for estimated traffic is the year of decision to build. The forecasting year is the first full year of operations (Flyvbjerg, 2005b).

4 For each of twelve urban rail projects, data exist for both cost overrun and traffic shortfall. For these projects average cost overrun is 40.3 percent; average traffic shortfall is 47.8 percent.


6 It has been verified with Daniel Kahneman and Dan Lovallo that this is indeed the first instance of reference class forecasting in practice. Personal communication, author’s archives.

7 The closest thing to an outside view in large infrastructure forecasting that the study uncovered is Gordon and Wilson’s (1984) use of regression analysis on an international cross section of light-rail projects to forecast patronage in a number of light-rail schemes in North America.

8 The lower limit of a one-third share of private risk capital for such capital to effectively influence accountability is based on practical experience. See more in Flyvbjerg, Bruzelius, and Rothengatter (2003: 120-123).
Oversigt over afhandlinger, som indgår i doktorafhandlingen "Megaproject Policy and Planning: Problems, Causes, Cures"

Afhandlingerne er opstillet i kronologisk orden. Der foreligger forfattererklæringer for alle gruppearbejder.


Megaproject Policy and Planning:
Problems, Causes, Cures

Megaprodukters politik og planlægning:
Problemer, årsager, løsninger

Sammenfattende redegørelse for
doktoraftandling af Bent Flyvbjerg