FLEXIBILITY IN DESIGN

(Working title, subject to change)

by

Richard de Neufville
Engineering Systems Division
Massachusetts Institute of Technology

and

Stefan Scholtes
Judge Business School and School of Engineering
University of Cambridge

PREFACE

This book focuses on the challenge of creating best value in large-scale, long-lasting projects through flexible engineering design. This best value has two components, first, the immediate value of the initial performance and, second, the long-term value associated with making the system adaptable to changing circumstances. An uncertain future provides a range of opportunities and risks. We can best deal with these eventualities; we can maximize our expected value, if we engineer flexibility into the project from the start.

Much of today's design practice focuses on "design for today", the achievement of acceptable initial performance in the short-run. This is not surprising. Government and public sector organizations normally invest in large-scale systems after they sense a compelling gap in today's service provision. Companies invest to exploit an opportunities that they have identified today. Closing an evident gap or reaping immediate profits appeals to users, politicians, chief executives, shareholders, to all involved.

However, any large-scale investment creates a value in two ways: short-term payoffs and complementary benefits in providing options for long-term adaptability. Unfortunately, the longer-term future is not clear, and long-term benefits are difficult to assess. Short-term, myopic planning thus tends to disregard long-term consequences. However, design that does not account for range of possibilities that may occur over a long lifetime risks leaving significant value untapped, risks underperforming overall. We argue for a better balance of the short and long-term sources of value in engineering design.

This book helps developers of major projects create value by using the power of design flexibility to exploit uncertainties in technological systems. You have the opportunity to increase the expected value of your projects significantly by cleverly designing projects to manage risks. Flexible design is key to success, as this book illustrates throughout. Designs that you can adapt to new circumstances enable you to avoid downside risks and exploit opportunities. You can thus use flexible design to improve your ability to manage your financial and social risks and opportunities. Technical professionals who can plan and execute a project to adapt to new circumstances can substantially increase the value obtained. The project team can use the design of projects to promote the needs to time, phase and diversify investments strategically.

This book is for all current and future leaders of the development, operation, and use of large-scale, long-lasting engineering systems. Your current or prospective responsibilities may include, but are not limited to, projects implementing:

- <u>Communication networks</u>: fiber optic cables, cellular devices, and fleets of satellites;
- <u>Energy production, transmission and distribution</u>: thermal and nuclear generators, hydroelectric plants, wind farms and other renewable sources;

- Manufacturing: for the production of aircraft, automobiles, computers, and other products;
- Real Estate: residential and commercial high-rises, hospitals, schools;
- Resource Extraction: oil exploitation and refining, mining and smelting;
- Transport: airports, highways, metro lines, high-speed rail, ports, supply-chains; and
- Defense systems: aircraft, ships, and armaments of all kinds.

The common feature of these long-lasting engineering projects is that they are all subject to great uncertainties. In general, it is impossible to know future circumstances and needs ten, twenty or more years ahead. Moreover, technology changes rapidly and disrupts previous assumptions and forecasts. New technologies both create new opportunities -- and can make previous investments obsolete.

The book is for the entire project team, which may include many different kinds of professionals. You may be current or prospective:

- Designers: the engineers and architects who create the physical implementations;
- Financial Analysts: who estimate the value of different designs, and thus shape them; Clients: the owners, public officials, and program managers accountable for the projects;
- Investors and Lenders: the shareholders, banks, pension funds and others providing the capital for the investments;
- Managers: controlling the operation of the facilities as they evolve over their useful life;
- Users: who operate over the system, such as airlines benefiting from air traffic control facilities or the medical staff of a hospital; and
- Regulators: the authorities responsible for safeguarding the public interest in these projects.

You all share the common problem of adapting the system for optimal performance as its requirements and opportunities evolve unexpectedly over its useful life. The clever design that enables you to take advantage of new opportunities will prove fruitless unless the managers of the system understand the design and can organize to use it. Conversely, the best managers of the system may have little scope to cope with future circumstances if the designers have not configured the project with the flexibility to adapt. Thus, even though you participate in the development and operation of the system at different times, and may not deal with each other directly, you will benefit from a mutual understanding of the desirability of coordinated design and management of your system.

Organization of the Book

We have organized the book into three parts to suit the range of audiences interested in using flexibility to improve the value of complex engineering systems.

Part 1 provides a rapid perspective on why flexibility is necessary and how it delivers value. It provides a high-level orientation to the concepts and methods. It may be sufficient to senior leaders who want to understand the issues. It also provides a comprehensive perspective that motivates the detailed chapters that follow.

Part 2 presents the methods needed to identify, select and implement the kinds of flexibility that will provide the best value. This section is for designers and analysts who will want to justify and implement flexible design. It covers the range of necessary techniques: procedures to forecast and anticipate the range of uncertainties; methods to identify the most promising kinds of flexibility to use; tools for evaluating and choosing the best flexible designs; and ways to implement flexible designs successfully over the life of the project.

Appendices provide detailed back-up explanations of the analytic tools and concepts used to identify and justify flexibility in design. Readers may benefit from one or more of these sections, depending on their interests and needs. This section presents brief but comprehensive presentations of the mechanics of economic evaluation and discounted cash flows; the economic rationale for phased development; the mechanics of statistical analysis used in forecasting; the process of Monte Carlo simulation to explore complex scenarios; and the basic financial concepts of options analysis. Importantly, this section provides a detailed discussion of the Flaw of Averages, the conceptual pitfall that traps so many designs in underperformance.

About the Authors

Both authors have extensive practical experience in the development and use of flexibility in design in many fields. These include: Aviation, Aerospace and Defense systems; Energy production and distribution; Health care; Manufacturing; Infrastructure projects; Mining and Oil and Gas Production; Real Estate development; Telecommunications; Transportation; and Venture capital.

The major companies and agencies the authors have worked with include BP, Codelco (Chile), Eni (Italy), the Far East Organization (Singapore), Ford, General Motors, Greater Toronto Airport Authority, GMR Group (India), IBM; Kinhill (Australia), Lloyds TSB; McKinsey, MITRE, Pacific Consultants International (Japan and Asia), Phillips, Secretaria de Obras Publicas (Mexico), Shell, Singapore Government, UK Department of Trade and Industry, US Defense Department, and US Federal Aviation Administration.

Richard de Neufville is Professor of Engineering Systems and of Civil and Environmental Engineering at the Massachusetts Institute of Technology and holds concurrent visiting appointments at Harvard University, the Instituto Superior Técnico (Lisbon), and the University of Cambridge.

His career has been devoted to the development and implementation of systems analysis in engineering. At MIT, he teaches the School of Engineering course on “Engineering Systems Analysis for Design”. He has written five major textbooks in the area and has a worldwide reputation for his expertise in airport systems planning and design. He was Founding Chairman of the MIT Technology and Policy Program. He has received many prizes and awards for excellence in teaching, innovation in education, and his textbooks.

Stefan Scholtes is xxxxx Professor in Management Science at the University of Cambridge and is a faculty member of the Judge Business School and the Department of Engineering.

He is responsible for the course on "Risk Management and Real Options", a core course on the Technology Policy Programme at the University of Cambridge as well as an elective on various Masters programs throughout the University. He also delivers tailored executive courses in risk and opportunity management at the Judge Business School.

Acknowledgements

The Government of Portugal has provided extensive support for the preparation of this book though its support of the MIT-Portugal Program, a major collaborative effort to strengthen university research and education in engineering systems analysis and design.

Many colleagues have encouraged and collaborated in our efforts to develop and demonstrate the value of flexibility in design. Notable among these are David Geltner, Manuel Heitor, Bob Robinson, and Olivier de Weck. We have also appreciated and benefited from the network of colleagues who have critically reviewed and shaped our work, including Luis Abadie, Gregory Baecher, Chris Caplice, José Chamorro, Gail Dahlstrom, Michel Haigh, Paulien Herder, Qi Hommes, Vassilios Kazakides, Afonso Lopes, Ali Mostashari, Joaquim da Silva, Eun Suk Suh, and Joseph Sussman. The challenging support of doctoral and post-doctoral students has been invaluable – our thanks go especially to Jason Bartolomei, Michel-Alexandre Cardin, Markus Harder, Rania Hassan, Konstantine Kalligeros, Yun Shin Lee, Jijun Lin, Katherine Steel, Tao Wang, and Yingxia Yang.

CHAPTER 1
INTRODUCTION AND EXECUTIVE SUMMARY

"We don't even know what skills may be needed in the years ahead. That is why we must train our young people in the fundamental fields of knowledge, and equip them to understand and cope with change. That is why we must give them the critical qualities of mind and durable qualities of character that will serve them in circumstances we cannot now even predict."

John Gardner (1961)

The future is uncertain

Technological systems can quickly become obsolete. New developments continually arise to displace established technologies. What was state-of-the art yesterday may be out-of-date tomorrow. We see this in our own lives. Consider the distribution of music for example: in a few decades, it has gone from vinyl records, to tapes, to CDs, to downloading tunes wirelessly onto miniature portable devices.

What happens to consumers also happens to large industries. The recent development of global communications offers several examples of unexpected rapid change. Much to the surprise of the developers of the Iridium and Globalstar satellite telephone systems, these were obsolete the moment they came into being -- ground-based cell phones had became universal (see Box 1.1). As further examples, wireless is substituting for landlines; satellite broadcasting is eliminating the need for local stations. Disruptive technologies pervade our lives.

Unexpected changes can create both gains and losses. System designers often equate uncertainties with risks -- and therefore with bad things. However, uncertainties can also create new opportunities. As with the internet, unexpected changes can create benefits that the original developers did not imagine. The future is as much about opportunities as risks. As the examples in Box 1.1 indicate, in thinking about uncertainties we should not just worry about downside risks – we need to keep upside potential in mind.

Box 1.1 about here

New technology affects the value of investments both directly -- and indirectly by the way it changes patterns of demand. Advances may have complicated, unanticipated ripple effects. Improved health care, for example, has increased life expectancy, which in turn has contributed to a greater population of older patients with complex co-morbidities. In general, the ultimate impacts of technological developments are complex and uncertain.

The potential benefits of any venture also depend on the vagaries of markets and many other factors. A copper mine may be lucrative if the price of copper is high – but not worthwhile if demand changes and prices drop. The benefits of any process also depend on its productivity;

the skill, experience and commitment of the staff; the success of marketing campaigns and the speed of diffusion of use; etc.

As this chapter shows, the bottom line is that we cannot count on accurately forecasting the long-term benefits and costs of technological systems. In general, the future value of these investments is highly uncertain. This is the reality that confronts designers, analysts, clients, investors, managers, users and regulators.

Standard methods are inadequate

Unfortunately, our methods of designing do not deal with the reality of rapid change. Standard practice proceeds from a set of deterministic objectives and constraints that define what the designers must accomplish. These mandates go by various names: systems engineers think of them as “requirements”, architects refer to “programs”, property developers and others think in terms of “master plans”. By whatever name, these restrictions channel designers toward a fixed, static view of the problem. In the case of the Iridium communications satellites, for example, the designers sized the fleet for worldwide use assuming 1 million customers in the first year of operation; they made no provision for the possibility of far fewer customers or a narrower service area. Likewise, in the extractive industries it is usual to base design on an assumed long-term price of the commodity, despite everyone’s experience that the prices of raw materials fluctuate widely. In practice, we “design to specification” when we should “design for variation”.

Our standard procedures for selecting designs likewise do not generally deal with the possibility of change. The standard methods for ranking possible choices refer to the “cash flow” of an investment, that is, to the stream of benefits and costs in each period of the project that would occur if the conditions assumed were to exist. In practice, the evaluation process usually discounts this unique flow and brings it back to a reference time to create measures such as the Net Present Value (NPV), the Internal Rate of Return (IRR) or the Benefit/Cost Ratio (see Appendix B for details). None of these approaches recognizes two routine features of large projects:

- The assumed conditions, such as demand and price, change as the project is nearing completion, as discussed earlier; and
- Management might – as it generally does – eventually decide to change the system in response to new circumstances.

Therefore, the initial business case analysis used to select design solutions often falls apart later on in the project. Consequently, the path that the project has chosen may be less than optimal.

The standard methods do routinely explore how new circumstances might change future benefits and costs. This is good, but standard analysis does not go far enough. Analysts calculate how different important factors – such as prices, market share, and rate of innovation – affect the cash flows and overall value of the projects configured to satisfy stated requirements. The difficulty is that this analysis of the sensitivity of a fixed design to alternate scenarios leaves out a crucial reality: the owners and operators of a project will alter the design in line with new realities. They may cut their losses by exiting from a project. They may increase their profits by taking advantage of new opportunities. They will in any case actively respond to new circumstances rather than submitting to them passively, as standard evaluation procedures assume.[1]

People design projects based on a limited set of assumptions and then think about uncertainties. They do not design with the uncertainties in mind. As the examples in Box 1.2 indicate, system designers do not generally explore how changes in the specifications and market factors would change the design itself. In short, our usual design and evaluation procedures focus on an unrealistically narrow description of the possibilities.

Box 1.2 about here

A great mismatch thus exists between what actually happens to a project over its existence, as the system owners attempt to maximize the value extracted from the project, and our standard ways of valuing and choosing between alternative designs. The reality is that future benefits and costs are uncertain, that future adaptation to evolving circumstances is commonplace, and that designers face a broad range of possible circumstances. Yet the standard valuation procedures assume that designers can design adequately around a single concept of the future benefits and costs – and that users will not deviate from a plan.

This mismatch matters. In general, a great difference exists between the value associated with the reality of many possible futures, and the value calculated on the assumption of a single future. This gap exists because equal variations up or down in the imagined futures do not translate into equal variations in values. All the possible futures do not happily average out to the most likely cash flow. To assume that the average cash flow generates an average value of a project is to make a fundamental mistake: to be a victim of the “flaw of averages” (see Box 1.3 and Appendix A). Designers absolutely need to deal with the broad range of possible futures.

Box 1.3 about here

Flexibility in design adds value

Once we recognize that the future is uncertain, the intelligent thing to do is to prepare for the various possibilities. For example, if there is a chance of rain when we go out, the commonsense approach is not to ignore this eventuality, but to prepare for it in some way, such as by bringing along an umbrella. “Be prepared” is a long-standing motto that applies not only to our personal life, but also to our professional life.

As both theory and many case studies demonstrate, creating designs that are “prepared” for future possibilities can add great value to a project. Flexible designs enable the owners, the managers or the operators to adjust the design to new circumstances. When the future turns out to be unfavorable, this can permit them to avoid bad consequences. When the future offers new opportunities however, flexibility in the design will enable them to take advantage and benefit from those possibilities (see Box 1.4).

Box 1.4 about here

Flexibility in design can easily lead to significant improvements in the benefits that you can expect to achieve. Thus, the case studies reported in the book show increases in expected value ranging up to nearly 80%. Flexibility provides a two-fold advantage: it limits possible losses and it increases possible gains. Both actions increase the overall average value of a project. Even relatively small chances of making major gains or avoiding disastrous losses can cumulate to important gains on average expectations. For example, just a 1 in 10 chance of doubling profits could increase average overall value by 10%. For major projects costing billions, the combined value of flexibility in design can be worth hundreds of millions (see Box 1.5).

Box 1.5 about here

What this book does

The book will provide the leaders of major projects with what they need to know to create value in technological systems by using the power of flexibility to exploit uncertainties. It shows how they can:

- Recognize uncertainty by replacing usual point forecasts with projections of realistic ranges of possible future outcomes;
- Understand and communicate the value of flexibility in adding value to design;
- Estimate the specific value that flexibility contributes to their project; and
- Implement a development strategy that profitably exploits the advantages of flexibility.

The central message is that designing a system with the flexibility to adapt to future needs and opportunities greatly increases its long-term expected value, compared to standard traditional procedures for developing and implementing projects. The book demonstrates this point with a wide range of practical applications.

The book is also pragmatic. It shows how project leaders of technological infrastructure can achieve these extraordinary benefits by building flexibility into their design, so that they can avoid future downside risks and exploit upside opportunities. It provides a 4-step process for developing flexibility in design:

- Step 1: Recognize the major uncertainties the project is likely to encounter. This step thus identifies the kinds of situations flexibility in the system needs to face.
- Step 2: Identify the specific parts of the system that provide the kind of flexibility best suited to deal with the uncertainties recognized in Step 1.
- Step 3: Evaluate alternative flexible designs and select the best for incorporation into the project.

- Step 4: Plan for eventual implementation of the chosen flexibilities, both by arranging making arrangements with the stakeholders in the process, and then by monitoring the conditions that would indicate whether and when they should implement the design flexibility.

Box 1.6 illustrates the process.

Box 1.6 about here

Box 1.1

Technological surprise: the Iridium fleet of communication satellites

The case of the Iridium fleet of communication satellites illustrates the sensitivity of technological projects to rapid changes in context. This system was a superb technical development – but a miserable financial failure.

Iridium originally consisted of over 60 satellites that communicated with each other and any point on earth. It provided consumers with wireless telephone service from any location to any other, provided they took the 3-pound satellite phone outdoors. Motorola designed Iridium in the late 1980s and deployed it a decade later. By that time, it was commercially obsolete -- cell phone technology had swept the market.

Iridium went bankrupt and sold for $ 25 million dollars, about ½% of the $ 4 billion investment.[2]

Technological surprise: global positional system (GPS)

The United States military originally developed GPS to control long-range missiles. The heart of GPS is a fleet of satellites constantly beaming signals, like lighthouses in the sky. Receivers can automatically triangulate these beams to locate themselves very precisely. Such chips are now commonplace in civilian applications. Aircraft can position themselves accurately when no radar is available for example. Cell phones have them. Drivers and hikers use them to find their way in remote areas.

GPS has created tremendous opportunities and value in ways unsuspected by the original designers. Because they did not anticipate this tremendous commercial success, they did not build the original GPS with any capability to benefit from it – they did not incorporate any way to charge a fee for this service!

Box 1.2

Standard design based on fixed assumptions: oil platforms

Deep-water platforms for extracting oil and gas from sub-sea reservoirs, as in the Gulf of Mexico or offshore of Angola, commonly cost about $1 billion each.

Usual practice is to design the size and location of these platforms based upon an assumed price of oil, $40 per barrel for example.[3] (The price assumed is a closely guarded corporate secret, because of its importance in contract negotiations. It represents assumptions about longer-term prices and thus differs from immediate spot prices. It also varies by company and over time.) The norm is to mandate designers to design for a fixed price of oil, even though it varies widely and events have proven trends to be unreliable, as Figures 1.1 and 1.2 show.

[Figures 1.1 and 1.2 here]

The effect of assuming a fixed price is to ignore oil fields that would be profitable at higher prices. This means that when prices are in fact higher and secondary fields might have been worthwhile, the platforms do not have easy access to these valuable reservoirs. Their exploitation would require a completely new project, which might not be economically feasible. The owners thus miss opportunities that a flexible, easily adjustable design would have exploited.

Box 1.3

Flaw of averages: capacity constrained facility

A simple example illustrates the “Flaw of Averages”.[4] Consider a constrained facility, such as a 5 level parking garage with a total capacity of 500 spaces.

Suppose that the

- annual revenue per space used is $ 10,000;
- annualized cost of the facility is $ 3,000,000; and
- demand for spaces depends on whether or not developers open a new office building, and is equally likely to be either 350 spaces -- or 550 spaces.

Table 1.1 shows the possible revenues:

[Table 1.1 here]

The average profit is then (1000 – 200) / 2 = 400. However, the profit calculated by using the average demand of 450 spaces is ($3600 – $3000) = $600 !

The reason for the difference between the true expected value ($400), calculated using the actual outcomes and the false value ($600), estimated using average outcomes, is that the gains on the upside are limited by capacity constraints and do not compensate for the losses on the downside. You cannot park 550 cars if you only have 500 spaces!

Box 1.4

Flexible design: Tagus River Bridge

The Ponte de 25 Abril, the first bridge over the Tagus River at Lisbon, offers a good example of flexibility in design. The Salazar dictatorship inaugurated it in 1966 with a single deck for automobile traffic, but with the strength to add on a second deck at some future time. Moreover, they also built a railroad station under the toll plaza, to minimize disruption in case Portugal ever decided to build rail connections.

A generation later, Portugal was a democratic member of the European Union, which allocated funds to develop commuter rail services throughout the region, and in 1999 the bridge received a second deck that carried these lines.[5]

When first built, designers recognized that the ultimate capacity of the bridge could be larger. Instead of trying to anticipate specific future requirements, they built for immediate use, with the flexibility to develop in many ways. In any case, if the original designers had tried to define future requirements, they could hardly have imagined the overthrow of the dictatorship and the creation of the European Union.

The flexible design of the bridge saved money by not building too early or building more highway capacity that might have been inappropriate. It also enabled Portugal to take advantage of the support of the European Union to create rail traffic across the river.

Box 1.5

Flexibility leads to major gains: satellite fleet

A detailed analysis of alternative ways to deploy geostationary satellites over different regions showed that a flexible system design, that enabled the system operators to reposition the satellites as demand for broadcast services changed, greatly outperformed the system "optimized" for the specified "most likely" pattern of demand.[6]

As Table 1.2 shows, the flexible design increases the overall expected value. Instead of launching the final fleet right away, it initially launches a smaller fleet – this simultaneously reduces the amount initial capital expenditure, thus the amount at risk and possible losses. The flexible design then deploys the second module sized and located according to actual need, thereby obtaining a maximum value if the demand exceeds original anticipations.

[Table 1.2 here]

Box 1.6

Application of 4-step process: high-rise building

Many developers have used "vertical flexibility" in the design of their buildings. The development of the Health Care Service Corporation building in Chicago illustrates the process. [7]The original design for this building has the strength, the space for elevator shafts and stairs, and the planning permissions to add 24 more stories to the original 30-story skyscraper built in the 1990s.

The 4-Step Process proceeds as follows:

1. The architects/developers recognize client's uncertainty about the amount of space needed, for example because it is not possible to be sure about either the long-term growth or how zoning regulations might change to allow them greater height.
2. They identify that vertical flexibility is the only realistic possibility, because they cannot increase the area available for development.
3. They explore numerous design alternatives, involving different numbers of floors for the first and later possible additions; estimate how each might perform under the range of future scenarios; and choose the arrangement that provides the overall best set of metrics.
4. They plan for implementation by making arrangements with the multiple stakeholders involved in the execution of the vertical flexibility, and monitoring developments to determine when (and if) they should use their design flexibility. They obtain planning permission for expansion, operational support from the tenants, financial commitments from the bankers, etc. They then keep track of their needs for additional space. The owners inaugurated the additional phase in 2010.

[Figure 1.3 here]

Table 1.1: Spreadsheet for parking garage example.

Spaces		Money, millions		
Demand	**Used**	**Revenues**	**Costs**	**Profits**
350	350	2.8	3.0	- 0.2
550	500	4.0	3.0	1.0
450	450	3.6	3.0	0.6

Table 1.2: Comparison of value of “optimized” and flexible designs for a satellite fleet. Note that the design “optimized” for a single forecast performs poorly on average across the range of possible scenarios. (Source: Hassan et al.)

Design	Present Value, Millions			
	Expected	**Maximum**	**Minimum**	**Fixed Cost**
“Optimized”	49.9	192	- 162	- 393
Flexible	95.8	193	68	275
Which Better?	**Flexible**	**Flexible**	**Flexible**	**Flexible**

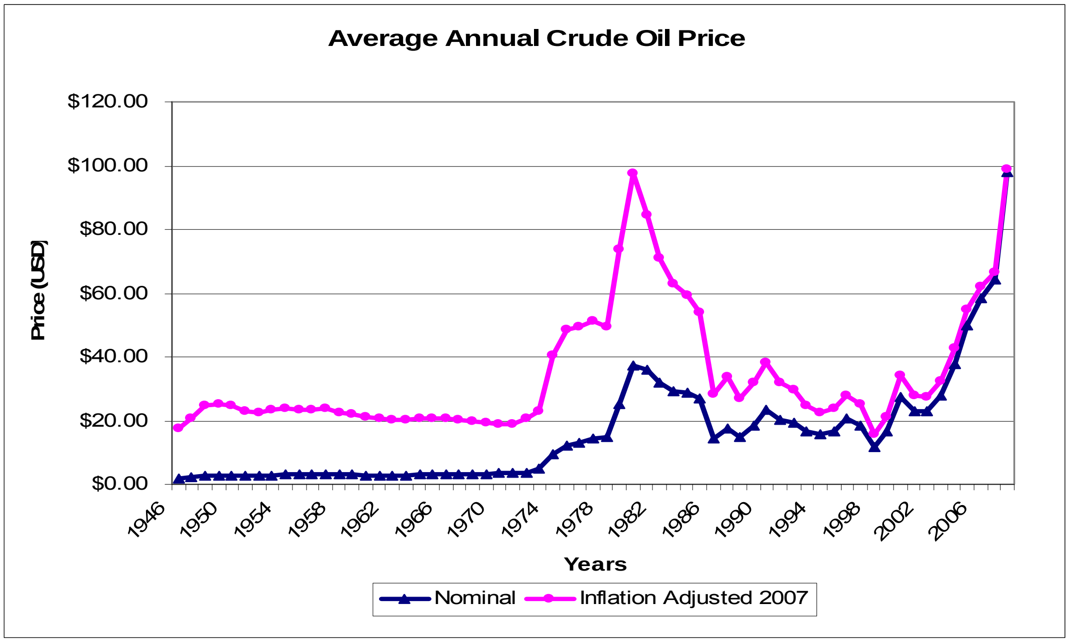


Figure 1.1: Historical prices of crude oil, both in nominal and constant value dollars. Notice how the 'trends' have frequently changed direction substantially: constant prices until 1973, a sharp run-up over the next decade, followed by a decade of dropping prices and the more recent reversal.

(Source: xxxxxxx [Need to update graph, of course.]

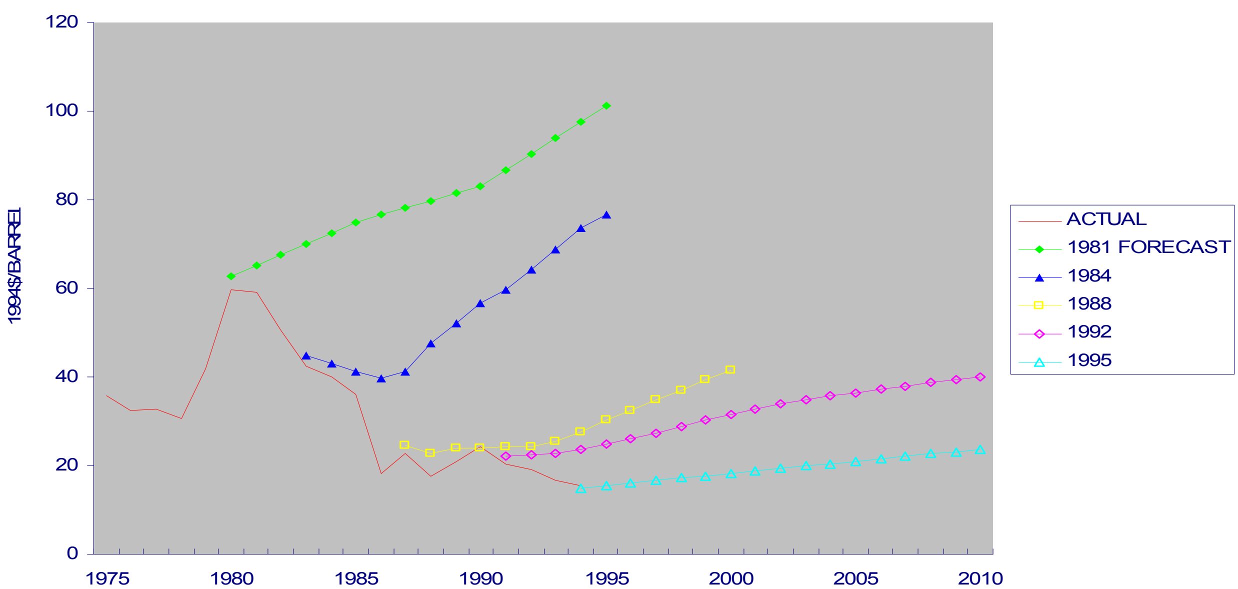


Figure 1.2: Forecasts of oil prices compared to actual prices. Notice how expert estimates failed to anticipate reality, even in the short run.

(Source: US Department of Energy, compiled by M. Lynch)

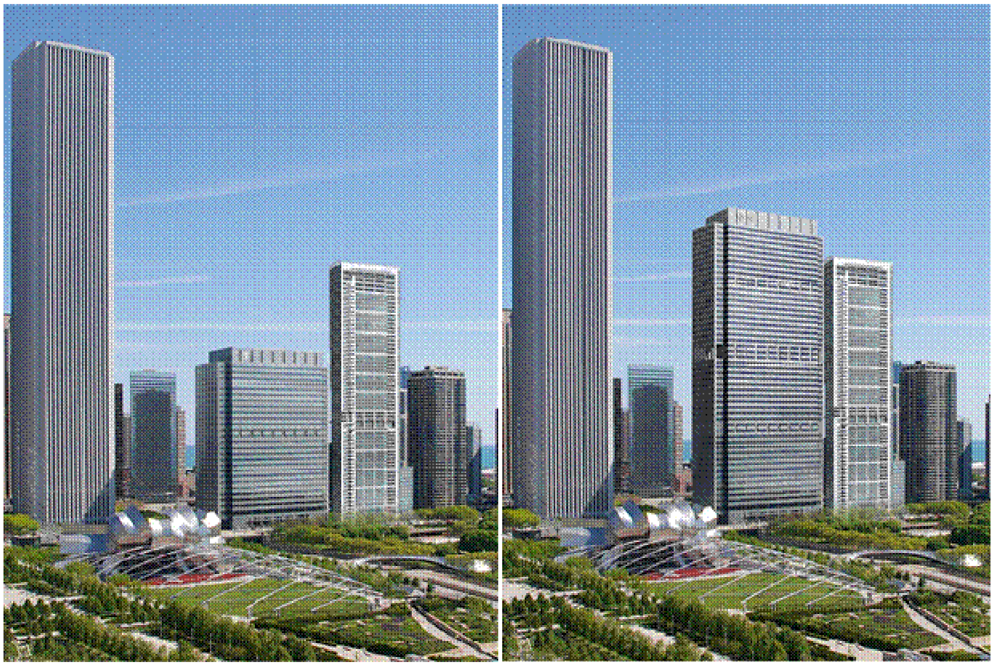


Figure 1.3: Vertical expansion of Health Care Service Corporation Building in Chicago in center of image. Phase 1 (left) and Phase 2 (right).
(Source: Goettsch Partners release to Wittels and Pearson, 2008).

[1] Inventory theory falls into this trap, too. Standard practice is to pick a level of service (percent of acceptable stock out) and then determine how much inventory to keep on hand in order to meet it. In practice, managers do more than just rely on safety stock – they expedite goods, cross ship them from a different location, etc. Inventory theory can therefore lead to excessively high inventory levels that do not make sense in practice.

[2] de Weck et al (2004) discuss in detail how the Iridium satellite deployment could have been better managed.

[3] Babajide et al (2009) discuss the standard design for oil field development and its alternatives.

[4] Savage coined the "flaw of averages" term. He describes the concept in detail in Savage (2009). Yang (2009) provides an extended example for the case of capacity planning in the auto industry. Appendix A discusses the mathematics.

[5] Gessner and Jardim (1998) provide technical details on this double-decking process.

[6] Hassan et al (2005) describe the analysis for the satellite example.

[7] Guma et al (2009) describe the case. Guma (2008) and Wittels and Pearson (2008) provide more details.

CHAPTER 2
RECOGNITION OF UNCERTAINTY

"In the early 1980s the consultants McKinsey and Company were hired by ATT to forecast the growth in the mobile market until the end of the millennium. They projected a world market of 900,000. Today [in 1999] 900,000 handsets are sold every 3 days."

A. Wooldridge (1999)

This chapter indicates how we might best anticipate the future for long-term technological systems. Modesty is the best policy. We need to recognize the limits to human foresight. We need to recognize that forecasts are "always wrong". People cannot realistically expect to get an accurate long-term forecast of what will happen. We might exceptionally be able in ten or twenty years to look back and see that our long-term forecast was accurate, but few are ever so lucky. We need to recognize that our future is inevitably uncertain. Therefore, we need to look at the ranges of possible futures, and to design our projects to deal effectively with these scenarios.

The success of our designs lies in how well they provide good service to customers and the public over time. To the extent we create systems that efficiently continue to fulfill actual needs, in the right place and time, we will be successful. If the systems do not provide needed services conveniently and in good time, our creations will have little value. Successful design thus depends upon anticipating what is needed, where it is needed and when it is needed. Success depends on the way we understand what the future might bring.

Unfortunately, it is almost impossible to predict exactly what the future will bring over the long term, over the life of our systems. Experience demonstrates, again and again, that specific forecasts turn out to be wrong, to be far from what actually happens. Moreover, because so much about the future is unpredictable, we cannot hope ever to achieve good long-term forecasts reliably. We need to recognize the inevitable uncertainty of any prediction. We need to give up on the idea of trying to develop accurate forecasts of long-term futures – this is an impossible task.

We need to adopt a new forecasting paradigm. We need to focus on understanding the range of circumstances that may occur. We need to consider the distribution of possible futures is often asymmetric, having limits in one side and being almost none on the other. For example, the price for a commodity is not going to drop below zero, but may become very high. Similarly, the maximum demand a plant might serve might be limited by its capacity, yet the minimum could vanish. We need to do this to appreciate the context in which our systems will function, to be aware of the risks that may threaten them, and the opportunities that might occur. This focus on the likely range of possibilities is fundamentally different from the standard approach to forecasting, which looks for the right number, for a single point forecast.

Uncertainty matters

Forecast uncertainty matters. If we do not take the range of possible outcomes into account from the start, we are almost certain to get wrong results from our economic appraisals, and to select projects incorrectly. Why is this? The intuitively obvious thought is that we could base our designs and evaluations on the most likely or average forecasts, and let the under predictions balance out the over predictions -- Why is it not right? These are good and important questions. The answer in a nutshell is that this intuition misses an important issue. If you work on the basis of average forecasts, you fall into the trap of the "Flaw of Averages", already mentioned briefly in Chapter 1. When we do the mathematics correctly, we find that basing designs and evaluating projects using the most likely forecasts is almost certainly wrong. To design and choose projects correctly, we absolutely need to consider the range of possible outcomes.

The "flaw of averages" justifies the concern with the range of distributions of possible outcomes. Understanding the difficulties associated with this fallacy is thus central to dealing properly with risks and uncertainties. This chapter therefore starts with a section that develops the understanding of the "flaw of averages".

The bottom line for this chapter is that we need to develop our understanding of the range of possible futures. In this way, we can anticipate what might be needed, and provide for these possibilities from the start. This approach will enable us to adapt our designs to what actually is needed or desirable, and will consequently increase the value of our designs – which is the point of the book.

The flaw of averages[1]

The "flaw of averages" refers to the concept that it is not correct to calculate the average value of a project based on its performance under average conditions. This may seem peculiar. Why don't "average inputs" lead to the "average outcome"? This section explains the basic elements of this puzzle.

We start by noting that the "average outcome" of a project is the average of its performance under all the possible scenarios in which it might exist. This is what we really need to know. We then need to recognize that different scenarios can have quite different consequences for a project. A higher demand for a product may – or may not – compensate for a similarly sized lower demand. The higher demand might not increase revenues, for example, if the system does not have the capacity to meet the demand. Conversely, a lower demand might be disastrous if the resulting low revenues do not cover the cost of the system and lead to bankruptcy. The asymmetry between the results associated with differences in scenarios is important – it can even be crucial. We cannot appreciate this feature by looking at the average scenario. We can only correctly evaluate the real results of a system by looking at the range of the scenarios.

An example illustrates the phenomenon. Imagine a game in which you roll a die and will receive $3 if the resulting number is 3 or more -- and nothing otherwise. Assuming that all the possibilities from 1 to 6 are equally likely, the average number showing will be 3.5. If you calculate your payoff based on this average number, you obtain a payoff of $3. However, the average payoff is only $2! This is because there is a 1 in 3 chance of rolling a 1 or a 2, which leads to a payoff of $0 instead of the $3 calculated based on the average number on the dice. This loss in the "downside scenario" is not counterbalanced by additional gains in the equally likely "upside scenario" of rolling a 5 or a 6, in which case you still only get $3. You lose in the downside but do not gain in the upside scenario.

This example is widely replicated in industry whenever it has to plan production capacity. Auto manufacturers, for instance, typically have designed plant sizes based on their most likely forecasts of sales of new models. If the model is a success and demand is greater than expected, they can expand capacity to a limited degree (and substantial cost) by working overtime or extra shifts. However, if demand is poor, then they definitely have big losses. (See Box 2.1 for practical application to the design of a xxxxx new example, Stefan to develop.)

[Box 2.1 about here]

The flaw of averages effect is easily understood – and remarkably often neglected in planning. Upsides and downsides of input uncertainties (numbers on the dice, demand for cars, use of parking spaces) will nicely balance out around their average, but these upsides and downsides will often have asymmetric effects on the measure of value for a system performance (such as its net present value). We simply cannot expect system performance to balance out around the performance calculated based on average scenarios.

The need for forecasts

Forecasts are fundamental to design. We plan and implement systems around anticipated future demands and opportunities. Our estimates of what we will need shape what we build, where we locate it, and when we implement it. These projections focus our thinking and determine the major characteristics of any design.

Meanwhile, the value of any design derives from how well it matches the actual realizations of future demands and opportunities. A system has value to the extent that it matches ongoing needs, at the right time and place. A "bridge to nowhere" may be technically brilliant, structurally elegant, and constructed efficiently and safely. However, if it leads nowhere, serves little traffic, and thus does not fulfill a meaningful function – it has little value. The development of the Iridium system to provide worldwide telephony illustrates the point. The deployment of this fleet of satellites that could talk both to each other and to any point on the planet was an astounding technological achievement. However, the system was a financial disaster. It was a $4 billion dollar investment sold for about $25 million or only about ½ of 1 percent of its cost. Overall, Iridium was not a successful design. It did not provide good value for the money spent.[2] Good design provides good value in terms of fulfilling actual demands.

Forecasts thus determine the value of a design. To be precise, our estimates of value depend on how well our projections of the range of future needs and opportunities match those that we eventually encounter. If we manage to anticipate correctly the possible scenarios that might occur, then our design can be successful. If our estimates are misguided, then we are likely to fail – either by exposing the project to possible risks, or by missing possible opportunities.

The first question for project leaders and designers must then be: how do we get the right kind of estimates of the future? The first part of the answer is that we must change our expectations about prediction. Current practice does not serve us well: it tends to develop forecasts that are unjustifiably precise. Most importantly, long-term forecasts are reliably wrong. We need to understand that in addition to trying to anticipate most likely futures, we must recognize the large uncertainties around these estimates. We need to be modest about our ability to predict. We need to recognize uncertainty.

Standard practice

Obviously, it is desirable to know what to design for. Consequently, the development of large-scale long-term systems always involves great efforts to forecast the future or, equivalently, to specify the requirements for future developments. For commercial or civilian projects, such as the construction of communication networks or oil platforms, these efforts are largely directed toward determining the context in which the system will perform – for example, the level of traffic to be expected or the quantity of oil and gas in an oil field. For military projects, the emphasis is on the specifying the "requirements", that is, the specific functions that a future system must fulfill – which of course results from forecasts of what threats the system will have to face. Either way, traditional process directs the work toward defining the future needs.

Point forecasts. Standard practice leads to very precise predictions or specifications. The almost universal result from a forecasting process is a series of "point forecasts", of precise numbers of what will happen in various years. The annual aviation forecasts developed by the United States Federal Aviation Administration provide a good example of this. Table 2.1 presents an extract from a large number of their forecasts on national and local traffic, for various types of activities. Notice the precision – 9 significant figures on a forecast almost 18 years in the future!

[Table 2.1 here]

Highly precise estimates of long-term futures are commonplace for all kinds of activities. Here are some samples, taken from publicly available documents:[3]

- The International Energy Agency (200X)provides energy forecasts, indicating for example that "Demand for oil rises from 85 million barrels per day now to 106 mb/d in 2030" ;
- The US Energy Information Administration (200x)predicts that wind energy production in the United States will rise from 24.81 gigawatts in 2008 to 42.14 in 2030;
- Gartner Inc. (2005) forecast the future sales of cells phones in the USA as "By 2009 the basic smart phone category will grow by 44.9 percent to 180.2 million units, while enhanced smart phones will grow by 67.9 percent to 98.5 million units"

Such highly precise forecasts are not credible. The most skilled technicians, using sophisticated scientific instruments, operating in tightly controlled laboratories, would be fortunate to achieve such levels of accuracy in measuring known substances. Estimates of what might happen, in the messy real world, years ahead, cannot be anywhere near this good. If we are lucky, our predictions might be within a few percent of eventual reality – optimistically, the first two figures of the forecast might be correct.

Yet the common feature for almost all point forecasts is that they provide no indication of their uncertainty, let alone of the degree of confidence we will have in them. They do not suggest the real possibility – the almost certainty -- that actual situations may be very different from the forecasts. The point forecasts are thus misleading. In fact, as shown below, there are typically large gaps between the point forecasts and what actually occurs.

[Box 2.2 about here]

Specialized professionals. Another important feature of standard practice is the way it is done. Specialized professionals usually manage the process of forecasting. They typically work in separate agencies or companies, distinct from the designers or planners for the systems under consideration. For example:

- Economists and statisticians develop models to predict the future levels of air traffic.
- Market analysts use consumer surveys and focus groups to define future sales of automobiles.
- Geologists develop best estimates of the size of oil fields, which they hand over to the structural and other engineers designing oil platforms.

In short, there is a great professional distance between the producers of the forecasts and those that will use these figures for design. Physical and institutional distance between these groups furthermore reinforces this separation.

This separation of tasks conveniently shifts responsibilities and provides convenient excuses for design professionals. If their product does not meet the actual demands, they can feel absolved of responsibility. They can argue that they did what they were told to do, and cannot be responsible for the mistakes of the forecasters – who are conveniently in some other organization and have probably left the scene by the time it is discovered that the forecast made quite a bit earlier does not correspond to what has now occurred.

The gap between the producers and users of forecasts is dysfunctional from the perspective of achieving good project design. Each group of professionals works in its own world with its own criteria and expectations. The overall result is that designers use unrealistically precise estimates during the design process. Standard practice is thus frequently trapped into fixating onto misguided concepts.

The professional context drives forecasters toward reporting precise results, even though they can easily report ranges. From the outside, companies and agencies desire precise forecasts and often ignore ranges, when available. They routinely spend significant sums on forecasting exercises – millions and millions. These fees impel the analysts to provide detailed answers. Indeed, who would want to spend large amounts to get a report saying that the future market can only be known within plus or minus 50% of the forecast? From inside the profession, the tradition is to judge models according to statistical tests of how well their models of behavior match historical data -- which is the immediate evidence at hand. The forecasting exercise is widely seen as a precise mathematical analysis subject to accurate calculation. Both the professional mind-set and the demands of the managers buying the forecasts thus drive the forecasting process to produce detailed numbers and reassuring precision.

Complementarily, engineers want to work with specific numbers. Engineering is supposed to be analytic and precise; not fuzzy or ambiguous. Furthermore, however complex the system, it is much easier to design to some exact requirement than to a range of possibilities. Moreover, managers want assurance that their plans are on target. Together, engineers and managers reinforce the demand for detailed, exact forecasts.

Unfortunately, the combined focus on precise forecasts is misplaced. The results are not worthwhile, as experience demonstrates. Sadly, it is hard to break the reinforcing professional dynamics that drive the focus on precise predictions. However, we are better off dealing with reality than wishes. In this spirit, we need to examine carefully what we can reasonably hope for.

The standard forecast is "always wrong"

Forecasts for any year -- about the future demand for products or services, future prices, future technological performance -- rarely tell us what actually happens. The actual values or levels that exist in the forecast year are, as a rule, different from what most people predicted. After-the-fact comparisons routinely demonstrate a big gap between the reality and the forecast. It is in this sense that a good rule of thumb is: the forecast is "always wrong".[4]

Some forecasts by chance turn out to be correct, of course. When many forecasters make predictions, some projections will turn out to be closer to the true value than others, some may even turn out to have been quite accurate. When the Swedish Central Bank examined

52,000 forecasts made by 250 institutions for inflation and economic growth in the U.S., Japan, Great Britain, France, Italy and Sweden during the period 1991-2000, they found that no organization was ever fully right – some would predict inflation well, but not growth, and vice-versa. In any year, some organization is most accurate, but this is largely a matter of luck -- the winners change from year to year. There is no consistent ranking, no clear indication of which forecast will be right next time.

The general rule that forecasts for particular projects are “always wrong” is widely documented. Study after study has compared forecasts for a system to what actually happens. The results show wide differences between what forecasters anticipated and the reality 5, 10, or 20 years later. Three examples illustrate the kinds of discrepancies that are routine.

Consider first the estimation of oil reserves. One might think that the amount of oil in a reservoir might be knowable. After all, the deposit came into existence millions of years ago and is fixed. However, the quantity of interest is not the overall size of the field; it is the quantity we can expect to obtain. As for any mineral deposit, what we really want to know is the amount that our technology can extract economically at market prices. This amount depends on many unknown factors. For oil, the economically useful size of the field depends on the:

- Technology – which is constantly changing so that deposits which were once inaccessible are now available to us;
- Market price of the product – for example, the huge deposits of oil in the tar sands of Alberta can only be economically extracted if the price of oil is relatively high;
- Structure of the oil field – the extent to which it is fractured and transmits pressure that drives the oil and gas out;
- Quality of the oil – such as its viscosity (which affects its flow) and its sulfur content (which affects its value); and of course
- Skill of the operators in managing the field by suitably injecting water (to maintain pressure) and combining flow from different wells (to keep viscosity sufficiently low).

In short, forecasts of reserves must be full of uncertainties, and are in fact (see Box 2.3)

[Box 2.3 about here]

Consider next the matter of forecasting the cost of new systems. Collectively, the track record is poor. New high-technology systems are understandably uncertain. One might think, however, that we would be able to estimate costs reasonably accurately for long established systems such as railroads, metro systems and highways which we have been deploying for over a century. Such is not the case. Exhaustive studies of worldwide experience show that the cost estimates supporting design and investment decisions are routinely off by 50% or more. Figures 2.2 and 2.3 show the case for major highway and rail projects.[5]

[Figures 2.2 and 2.3 about here]

Apparent counter-examples of good predictions of future system costs are often not what they seem. The owners of Terminal 5 at London/Heathrow airport boasted that they delivered the project on time and budget – a great accomplishment due to innovative management. What they did not advertise is that the project's final cost of over £4 billion was well over twice the cost the developers originally anticipated when they decided to proceed with the project over a decade before it was opened. In any case, developers often meet their budget by cutting down on the design, so that they such comparisons often do not compare like with like.

Consider finally the matter of forecasting demand for the services of a system. This is especially hard to anticipate correctly. Demands for services result from a complex interaction of market forces, competition from other providers or other services, and from changeable personal preferences. Forecasting demand is especially difficult because we need not just anticipate the overall, aggregate demand, but its components. For example, a good estimate of the number of cars that will be sold is not good enough for an automobile manufacturer. The company also needs to know how many different kinds of vehicles it could sell, because it has to design and equip its factories differently to produce sedans, hybrids, minis and SUVs. Any company or organization delivering a variety of products or services must be concerned with forecasting each of these elements, and this is especially hard to do (see Box 2.4).

[Box 2.4 about here]

Requirements are also often wrong

"Requirements" provide an alternative way to specify what a design or project should look like. Forecasts define the opportunities or demands on a system -- for example, the size of an oil field or the loads on a commuter rail system. They guide the designer to a set of functions the client wants the prospective system to fulfill. "Requirements" directly specify what these functions and capabilities should be.

Governmental and other organizations routinely focus on "Requirements". They do so for areas where numerical forecasts might not be available or useful. The military routinely go to great efforts to specify the requirements for future systems. To do so, they will work with experts and consultants to define what future threats might exist, or what new capabilities their future strategies might require. Similarly, organizations in charge of air traffic control specify the performance standards for their equipment for monitoring and controlling aircraft.

Requirements frequently change. Although the name gives the impression that the elements of the requirements are necessary and absolute, in fact what developers once proclaimed to be "required" often turns out to be unneeded or no longer necessary. This happens when events, or perceptions of the situation, or the technology change what the system managers see as needed. The evolution in the way the military and other agencies have thought about unmanned aerial vehicles (UAV) illustrates how this occurs (see Box 2.5).

Overall, the point is that our estimates of what might be desirable and necessary are not reliable. They are uncertain. Understanding this fact and incorporating it into our design practice gives us the basis for creating the most realistic design for our systems.

[Box 2.5 about here]

Inescapable obstacles to accurate forecasting

Forecasts are not only "always wrong" as a matter of historical record, they are bound to remain so. This is because there are two inescapable obstacles to good forecasting. The first consists of "trend-breakers". These are the ever-changing set of surprises that disrupt our expectations. The second set consists of the inevitable ambiguities that must forever confuse our interpretation of the historical record.

Trend-breakers. These are events that disrupt the smooth continuation of what has been happening in the recent past. They come in all kinds of forms:

- Economic crises, as in 2008, but also in 2001, in 1991, and many other times;
- Political shifts, both peaceful (the creation of the North American Free Trade Agreement, NAFTA), and not (wars and revolutions);
- New technologies, such as the development of terrestrial cell phones that destroyed the potential market for satellite phones and led to the bankruptcy of Iridium;
- Discoveries, as illustrated by the doubling of the expected reserves in one of the oil fields;
- New market conditions, such as environmental regulations or the emergence of major new competitors; and so on.

Trend-breakers happen routinely. Any particular one may come more or less as a surprise; the record shows that we must expect such surprises. They regularly disrupt long-term forecasts and create new conditions that the designers and managers of new systems need to face. They shape large forces across the entire economy, and occur locally and regionally. They may affect an entire industrial sector, or only a product line. They are real – and a basic reason why any trend must be associated with considerable uncertainty.

What is the trend? Trends themselves are ambiguous. Given the historical ups and downs, either in overall demand for some services or in our perceptions of the quantity of oil in a reservoir, how do we establish a trend? This is not an issue of mathematics or graphics. Indeed, the procedures to develop a trend from any given set of data are well established and routine. The issue is: what set of data should we consider for establishing a trend?

The problem in establishing a trend is that the results depend on the data chosen, on the number of years considered. If one chooses a very short period, then the short-term effects may dominate and obscure a longer-term trend. For example, looking at the air traffic for Boston from

1999-2004, one could see an average drop in traffic of about 2% a year. This is a direct result of the burst of the dot.com bubble and the concurrent 2001 terrorist attacks. However, choosing a very long period, such as 50 years may cover several trends. Indeed, air traffic in the United States and Boston enjoyed a spurt of growth in the 1980s, following the national deregulation of airlines, the overall lowering of prices and increased competition. That spurt has since dissipated and we might best exclude it from any analysis of the current trends. Following that logic, the appropriate period for projecting might start somewhere in the 1990s. But when? Which years of data should we include in order to establish a trend?

The years selected for establishing a trend can easily change the long-term projection drastically. This is because slight changes in the pointer multiply into substantial differences when projected far. The example in the Box 2.6, based on data at a major airport, demonstrates how slight modifications in the range of historical data chosen can change the 10-year forecast by 20%. Trends are not obvious. They are in the eye of the beholder.

[Box 2.6 about here]

What drives the trend? We might reasonably ask: what drives the trends? What factors increase the future development? What might encourage users to want to use new communication services? Or hybrid cars? Or any of the products and services our systems might deliver?

The process of determining the drivers of future forecasts is also highly ambiguous. This is because it is rarely possible to separate out the factors that determine the change in anything in which we might be interested. As it turns out, most possible explanatory variables follow similar patterns, and it is almost impossible to unscramble their separate effects.

The pattern of change of most developments can often be reasonably described for a limited time by some annual rate of growth, call it r percent a year. This means that, starting from a current level L, the future level after T years is:

Future level, $L_T = L (1 + r)^T \sim L e^{rT}$

This pattern is called exponential growth. Moore's Law describing the doubling of computer capacity every two years is one example of this phenomenon. Exponential changes can sometimes last a long time; Moore's description of change held for over 4 decades since the 1960s. However, there are clearly limits to such patterns. In general, they last for some limited time. The sales of new products, for example, frequently grow exponentially during early stages, and then slow down as markets become saturated.

Exponential change generally applies not only to the trend of the factor that we want to forecast – but also frequently to several of the other factors that might drive or otherwise influence the factor of interest. For example, the factors influencing the increased use of cell phones might be the steady drop in price for equivalent functionality, the increase in population, changes in overall wealth -- all of which might reasonably experience exponential change, at least for a while.

Because an exponential process describes the rate of change of many possible drivers, we cannot untangle their effects. Technically, the potential drivers are likely to be highly correlated. The consequence of this high correlation is that it is not possible to have much – if any – statistical confidence in any of the effects of the drivers. Box 2.7 describes examples of this phenomenon. From a practical point of view, we simply cannot have much confidence in the results of analysis trying to determine the causes of most trends we would like to project.

[Box 2.7 about here]

Recommended procedure

*Recognize uncertainty***.** The starting point for good design is the fundamental reality that we cannot predict the future precisely. Our best thinking, our most thoughtful analyses cannot overcome the unpleasant facts that make forecasts unreliable. Interpreting the historical record is ambiguous at best. And even though trends may persist over the short-term, trend-breakers routinely occur over the long-term horizon for systems design.

We must recognize the great uncertainties that inevitably surround any forecasts. As the examples illustrate, the realistic range of possibilities is often as wide as a factor of two from a medium estimate for innovative, large-scale projects. The range of possibilities is large even for mature, well-established activities; witness the 10% deviation from forecast for air traffic for Boston after only 5 years. We should direct an important part of our forecasting effort toward developing a reasonable understanding of our uncertainties. What is the range of possibilities we may have to deal with over the life of the system?

We need to think carefully about the range of possibilities. Sometimes we can reasonably assume that their distribution is symmetric about the forecast of the most likely future, for example “plus or minus 20 percent”. However, the distribution is often significantly skewed. For example, the most likely long-term price of copper is probably in the range of US$ 1 per pound. Its lowest price will be well above zero, a drop of perhaps US$ 0.50 – but its highest price could well be as high as US$ 4, as has already happened. The distribution of copper prices is thus skewed from the most likely forecast – perhaps only minus 50%, but possibly plus 300%.

A good way to develop an understanding of the size and the shape of the range of uncertainties is to look at previous experience in the same or a similar area. The historical record for the product or service often gives a good view of the situation. Thus, a range for the price of copper is easily available from records of past prices. Figures 2.5 and 2.6 respectively show prices over the last 5 years for example, and back a century.[6] This kind of information gives a realistic perspective on the range of possibilities. For completely new situations, for which there is no historical record, an effective approach can be to look at situations that appear comparable. We might ask, for example, about the track record of predictions for new electronic technologies, when thinking of a new system in that field. We might look at previous experience predicting the size of oil reservoirs, when faced with a new site we have just begun to explore. In all cases, we should ask: what is the best information we can have about the size and range of uncertainties?

[Figures 2.5 and 2.6 here]

Where possible, we need to characterize the distribution of the uncertainties. We will be able to weigh the relative value of different designs more accurately to the extent we can estimate the probability of different outcomes. In some cases, we might be able to do this easily. For example, we can derive the occurrence of earthquakes, and other natural hazards unaffected by man, from historical records with some accuracy. More generally, we must use sophisticated statistical analyses to estimate the exact nature of the uncertainties (see Appendix D on Forecasting.)

Shift forecasting emphasis. To develop a realistic view of forecasts requires a double shift in the forecasting process. Most obviously, it is desirable to spend effort on the process of recognizing the uncertainties, defining their ranges, and estimating the probabilities of the outcomes. This implies a shift in emphasis, away from standard efforts to identify the most likely or best forecasts and point estimates, and toward spending more time and money thinking about the range of possibilities. Managers and users wanting the most realistic – and thus most useful – forecasts require the forecasters to identify and characterize the uncertainties.

The shift in emphasis involves a shift in mindset and expectations. To the extent that forecasters realistically accept the evidence that trend-breakers and other factors disrupt even the most sophisticated estimates of the future, they must also recognize that it is not useful to try to develop very detailed point forecasts. This will mean also that they can generally forego the very detailed kinds of analyses on which they currently spend so much effort. It is desirable not only to spend more effort on understanding the nature of the uncertainties, but also to spend less on trying to get best estimates of most likely forecasts.

Track developments. All the above implies that we need to track developments and readjust our estimates of the future accordingly. This is the logical consequence, once we recognize that forecasts done at any time are unlikely to be reliable due to trend-breakers and other changes in the environment. To develop designs that will be able to meet the current and future needs most effectively, system managers need to be on top of these needs. They cannot rely on forecasts made long ago to be the proper basis for design, even if some administrative process approved these forecasts. Official approval does not define reality!

To track developments most effectively, the forecasting process can identify leading indicators of change. These are factors that are most likely to signal changes in trends. For example, the developers of the Iridium satellite telephone system might have identified “sales of cell phones” as a leading indicator of the success of their prospective competition. If they had done so early enough, they might have redesigned their system and avoided much of its financial failure. System managers can then track these factors over time, to obtain early warning of trend-breakers and other developments that might affect the design. They should reallocate their forecasting efforts not just from point forecasts to understanding distributions, but also from a one-shot effort to a continuous process.

Take-away

Overall, system owners, designers and managers need to change their approach to forecasting. They should avoid asking for the "right forecast" or a narrow range of possible futures. If they request and pay for such a result, they will surely get it. However, it will be almost meaningless. Forecasts are "always wrong" in that the future that actually occurs routinely differs from what someone predicted.

Users of forecasts need to recognize the inevitable uncertainties. They should thus ask for a realistic understanding of the range of possible scenarios. Specifically, they can tailor their requests for forecasts to:

- De-emphasize the development of precise best estimates – such point estimates are demonstrably unreliable, generally very expensive, and thus a waste of money;
- Develop a realistic assessment of both the range of uncertainties around the forecasts and the possible asymmetry of any distribution;
- Estimate, to the extent possible, the probabilities associated with possible outcomes; and
- Track developments over time to keep abreast of the situation and adjust designs and plans accordingly.

Box 2.1

Valuing a XXXXXXX new example – Stefan to develop

The value of a Parking Garage depends upon the number of users. In many cases, these are uncertain, for example when the garage serves a shopping mall whose future success is unknown.

In this particular example, explained in detail in Chapter 3, the average actual system performance, measured by the average net present value of about $1.9M, is quite different from the system performance of $3.1 M obtained from the estimate of the average demand scenario. This deviation of the estimate from the performance of the fixed projection model is too stark to be a chance effect. The great difference between the estimate of the actual average value obtained by considering the various demand scenarios, and the number calculated using an average demand, is due to the capacity limitations of the garage. It cannot accept additional cars if the demand is high, so it cannot counterbalance the losses if demand is low. Thus, the actual results are much lower than anticipated by an analysis based on average values.

Box 2.2

The gap between designers and forecasters

One of the chief designers of the Iridium satellite system for global telephone service visited MIT after the financial collapse of the venture. In sharing his experience with the design, he indicated that one of his greatest difficulties was getting the correct forecast of future use, and therefore of fixing the specifications of the system.

What the designer wanted was a definite fix on the number of users. He wanted to use this number to optimize the size and capabilities of the individual satellites and of the fleet. The difficulty of getting this commitment from the marketing and forecasting team frustrated him.

He was of course asking the wrong question. No forecasters could precisely anticipate the actual demand for an untested service that would be deployed a decade later in competition with other rapidly evolving forms of technology – cell phones in particular. In retrospect, the forecast he used was about 20 times too high.

The right question for the designer would have been: "what range of usage might exist?" The answer to this question would have given him some idea about how he could stage the deployment of the system, to learn about the actual demand, and then to know how to shape the future development of the system to accommodate the actual needs.

Box 2.3

Variation in the estimate of oil reserves

Figure 2.1 represents experience with two oil fields in the North Sea, normalized around the size of the original estimate.[7] The P50 represents the median or most likely estimate at any time, and the P10 and P90 are the limits defining the range that the actual quantity expected to occur 80% of the time (the 80% confidence limits). Although one might imagine that the estimates would steadily converge on some value as more experience and information develops, it often happens that this experience reveals new finds or difficulties that trigger a "jump" in the estimates, as Figure 2.1 shows.

[Figure 2.1 here]

This figure is hard to interpret – needs to be redone

In the case on the left, the range of the estimate does eventually narrow considerably, but the most likely P50 estimate is less than half the original amount. In the case on the right, the range of estimates hardly narrows, and the most likely estimate nearly doubles. Notice moreover that over one year the best estimate rises dramatically; far exceeding what was considered the previous range of likely possibilities.

In this context, we note that design teams often focus on the original best estimates – the P50 values -- to design the platforms and wells for the field. For the 2 fields shown in the figure, this practice means that the resulting platform designs would be off by almost a factor of 2 -- about double the desirable size for the case on the left, and only half what would be needed for the case on the right. Although the geologists know from experience that estimates can jump over time, and although operators constantly struggle with production uncertainties, standard design practice does not account for this uncertainty.

Box 2.4

Difficulty in predicting demand and its components

Forecasts for Boston airport illustrate the difficulty in predicting traffic demand. In the case shown in Table 2.2, the discrepancy between the overall passenger forecast and reality was 10% over only 5 years. In retrospect, this was due to the burst of the dot.com economic bubble – one of those unexpected disruptive trend-breaking events that so frequently and routinely mess up forecasts.

The 10% discrepancy was not even across the board however. The number of aircraft operations dropped by twice as much (23%), as airlines cut out flights by small regional aircraft (42% off the forecast) and squeezed more people into their big jets. This kind of differential changes in demand over product categories is typical.

[Table 2.2 here]

Box 2.5

Changing requirements – unmanned aerial vehicles

Unmanned aerial vehicles (UAV) are aircraft that do not carry a pilot.[8] They range in size from that of model planes operated by hobbyists to much larger aircraft with wingspans of about 20 meters and costing over $8 million each. They differ from earlier drones in that ground operators stationed at great distances, often halfway around the world, can tightly control them – the US Air Force has been directing flights over the Middle East from a base in Nevada.

UAVs became practical with the development of miniature electronics and global communications. They are thus a product of the 1990s. They thus represent a new technology. Users are gradually understanding how they can best use them.

The "requirements" for UAVs have thus been changing dramatically as the users gain experience with them. Originally, the military imagined them to be surveillance aircraft that would loiter slowly over an area. The requirements were therefore for relatively slow vehicles (to make it easier to get good images) with modest payloads such as cameras and electronic sensors. More recently, the requirements for the same vehicles have shifted towards carrying heavy loads (such as bombs weighing a quarter ton) and greater speeds (to escape anti-aircraft fire). Concurrently civilian users, such as the Forest Service interested in spotting fires, have other requirements such as those associated with sharing airspace with commercial and other aircraft with people in them.

Box 2.6

Ten-year traffic projections

Figure 2.4 shows data for a major airport. As for a number of other major American airports, its traffic dropped in the late 1990s, in this case by about 20%. This was due to the increased competition from low-cost airlines at secondary airports, and then the burst of the dot.com bubble. The traffic then grew steadily until 2008, when another drop in traffic occurred.

The trend we extract from this history depends on the years selected. Suppose we wanted to project ahead 10 years from 2008, and thought it reasonable to look back 10 years for this purpose. If we consider the period 1998-2008, then we project ahead from about 15 to 18.5, as seen in the graph. However, as has been observed in the past, forecasters might assume that the dip in traffic in 2008 was a temporary dip and prefer to prolong the steady trend that seem to hold between 1998 and 2007. Projecting ahead from that period leads to a forecast of over 20. Another defensible view might be that the analysis might best not base a trend from a previous low to a recent high, and ought to include data from the earlier high. Following this logic to include all the data from 1996 to 2008, the resulting trend leads to a forecast of about 16.4. As the figures show, slightly different assumptions scatter the trends over a range of about plus or minus 10%. Trends are not obvious!

[Figure 2.4 here]

Box 2.7

Ambiguous analysis of causes of trends

A regular classroom exercise at MIT involves working with extensive historical data on traffic in Los Angeles. The purpose is to explore how the experts on the actual job arrived at their forecasts from this data. One phase of the exercise considers the expert analysis that selected the relative importance of population and income as drivers of traffic growth, based on which combinations matched the historical data most closely.

The class exercise invites the students to match the Los Angeles data with any other single factor. In recent years the factors that most closely matched the data, and in each case did so much better than the model used by the consultants for Los Angeles, were the:

- Divorces in France,
- Egg production in New Zealand, and
- Prison population of the State of Oregon!

Of course, this is all nonsense. That is the point of the exercise. By themselves, good statistical matches of annual data means very little, if anything, and it is unrealistic to use these procedures to develop and select useful forecasts.

Table 2.1 Forecast emplanements at United States airports.

Note assumed ability to forecast to a single passenger 20 years ahead.

Size of airport	Actual, 2006	Forecast, 2025
Large Hubs	507,958,226	892,531,074
Medium Hubs	146,257,477	250,288,322

Source: US Federal Aviation Administration (200X)

Table 2.2 Actual and forecast emplaned passengers and flight operations at Boston/Logan airport

Category	Actual, 1999	Forecast, 2004	Actual, 2004	Percent Under
Passengers	13,532,000	14,500,000	13,074,190	10
Operations				
Total	494,816	529,129	409,066	23
Big Jets	255,953	265,058	236,758	11
Regional	189,680	222,949	128,972	42
Cargo	11,267	13,180	12,100	7
General Avn.	37,916	27,942	31,236	12

Source: Massport and US Federal Aviation Administration (200x)

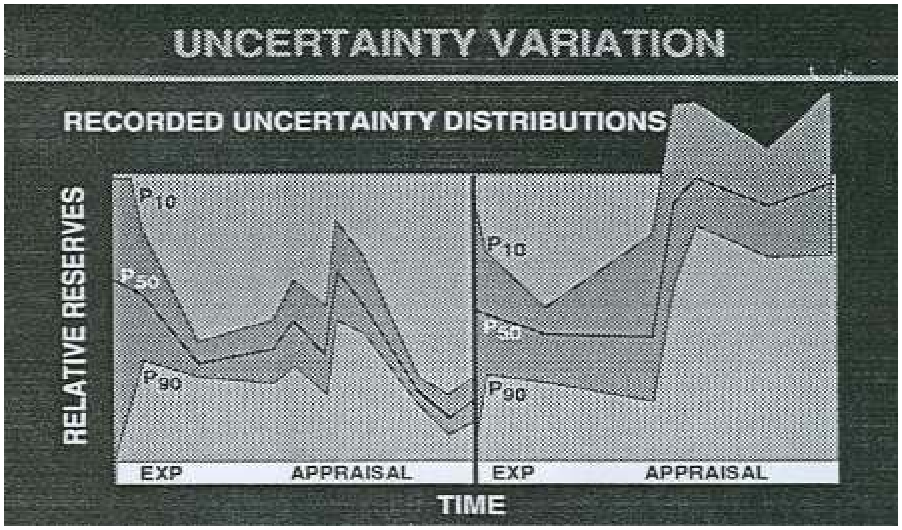


Figure 2.1: Example variation in best estimates of oil that can be commercially extracted from two fields (Source: Lin (2009) from BP sources).

This figure needs to reorganized

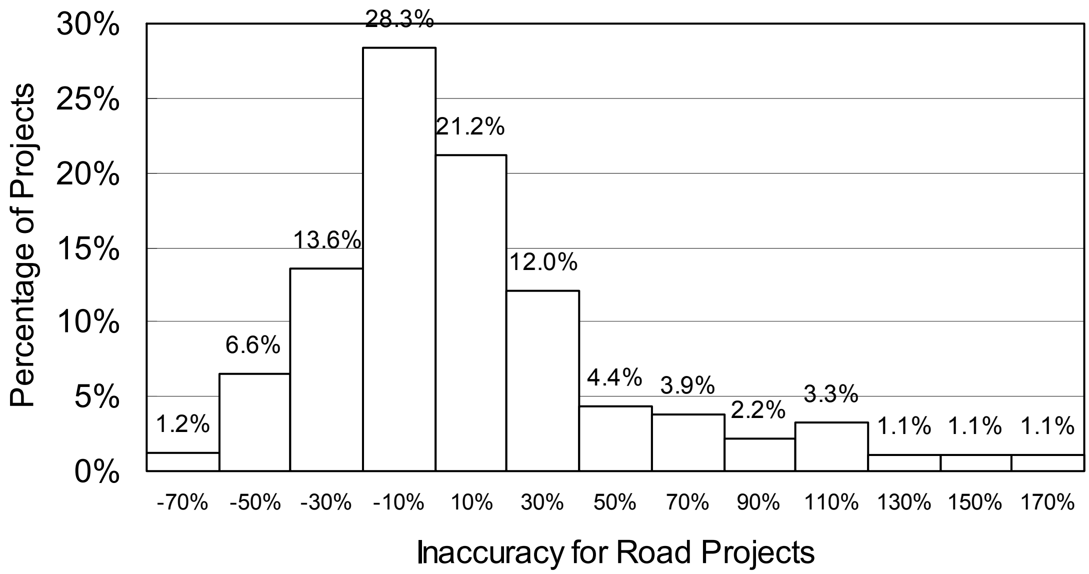


Figure 2.2: Discrepancies between the forecast and actual costs of road projects, adapted from Flyvbjerg et al (2005). **Need to specify units – cost or demand…redo 2.2 and 2.3 to be similar.**

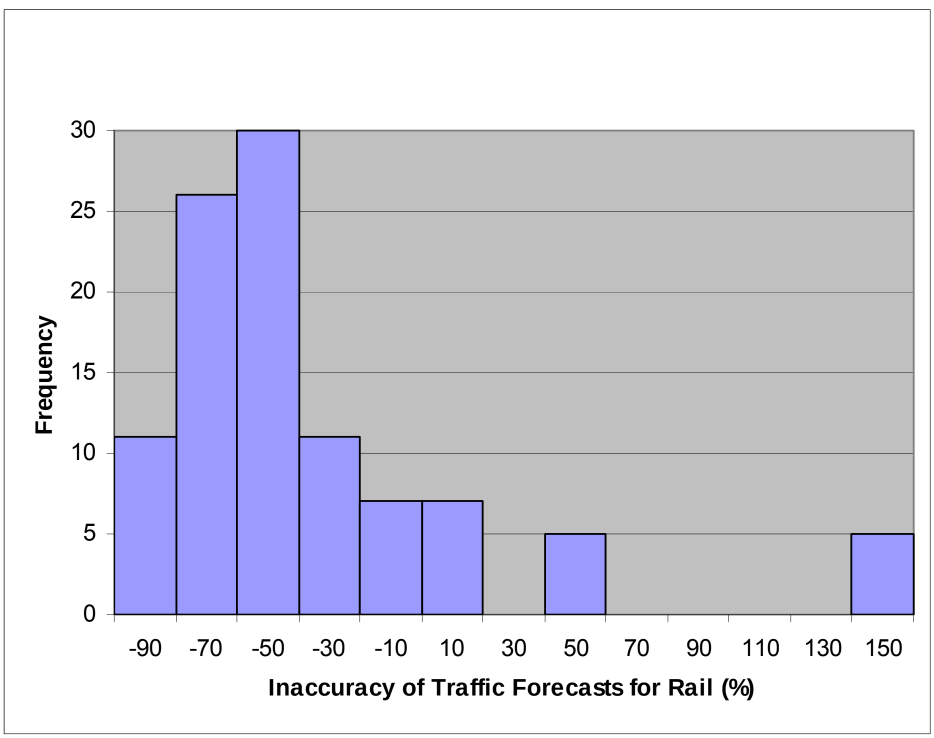


Figure 2.3: Discrepancies between the forecast and actual costs of rail projects, adapted from Flyvbjerg et al (2005).

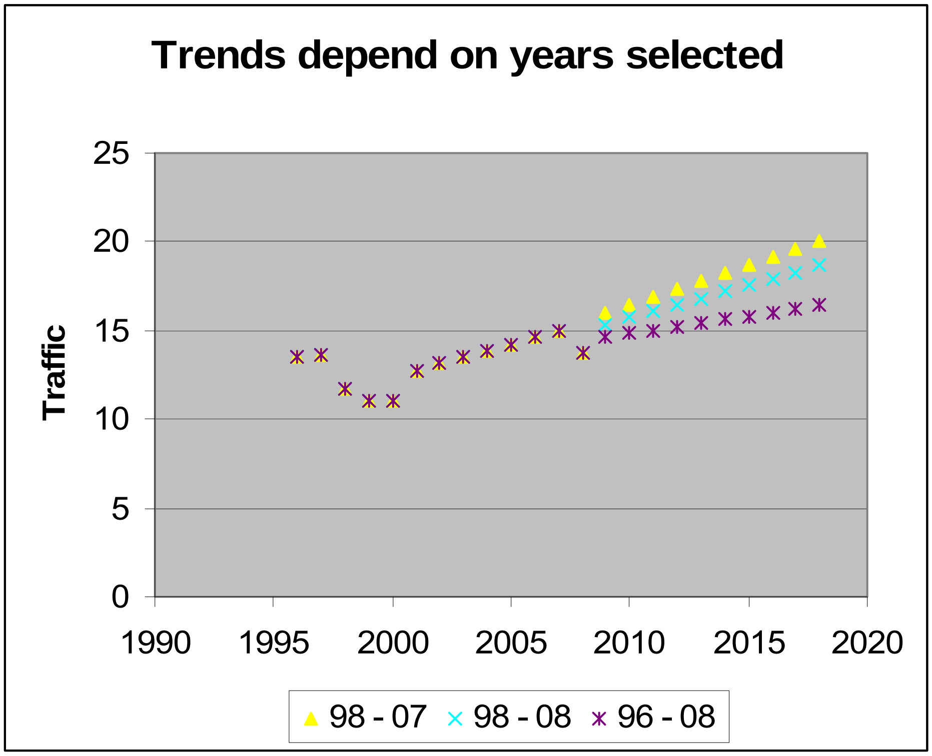


Figure 2.4: Traffic data for a major US airport

(Source: Based on data from the US Federal Aviation Administration (XXXX))

Redo to eliminate box at bottom and label lines explicitly in graph.

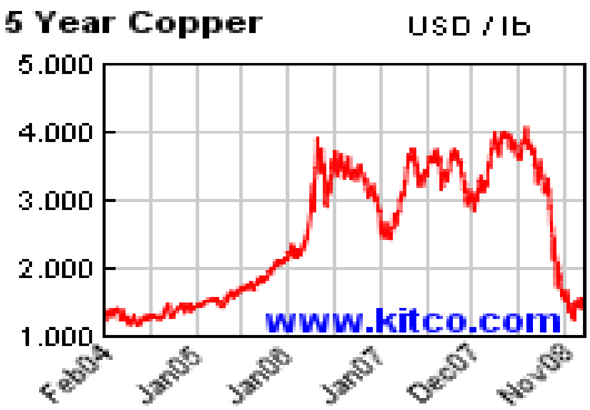


Figure 2.5: Range of copper prices, February 2004 – 2009 -- update
(Source: Kitco.com (2009))

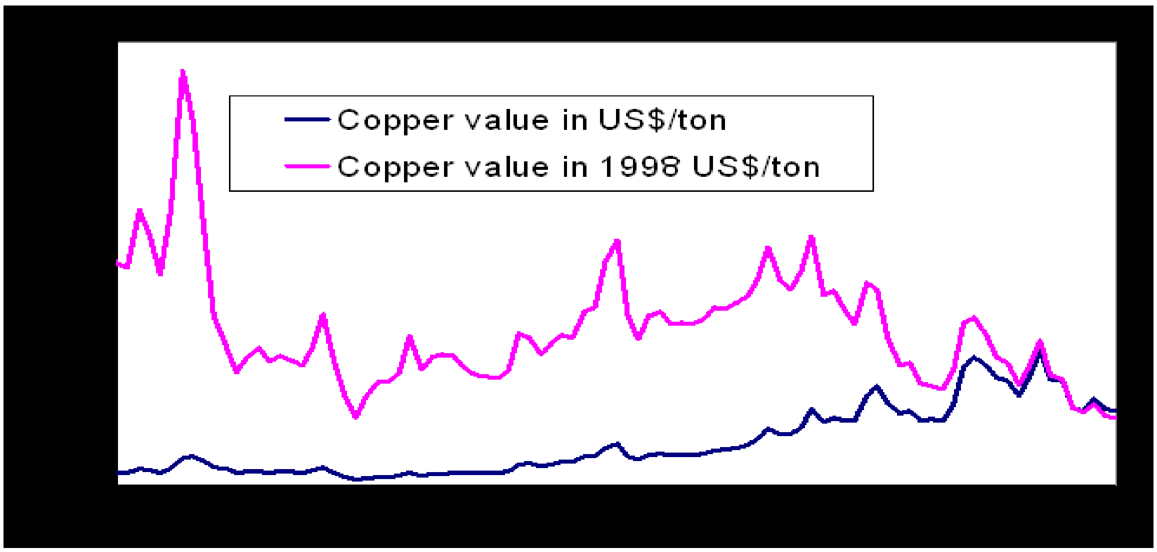


Figure 2.6: Range of copper prices, 1910 to 2003 update
Source

[1] Savage coined the "flaw of averages" term. He describes the concept in detail in Savage (2009). Appendix A discusses the mathematics.

[2] Chapter 1 discusses the Iridium case in Box 1.1.

[3] The forecasts cited are published regularly by these sources: International Energy Agency (IEA) *World Energy Outlook*, XXXXXX: (annual); US Energy Information Administration (EIA) XXXXXXXXXXXX: (2008); Gartner, Inc. XXXXXXXXXXX; and the US Federal Aviation Administration, "Table S-5, Terminal Area Forecast Summary, Fiscal Years 2007-2025", Washington, DC.

[4] Several comprehensive or meta studies of the inaccuracy of forecasts are available. Ascher (1978) provides a seminal text. Miller and Lessard (2001) give extensive examples of the inability of planners and designers to estimate uncertainties accurately.

[5] The data on variations between forecast and final results come from the extensive work Flyvbjerg and his team have presented in many places. See in particular Flyvbjerg et al (2005) and Flyvbjerg et al (2003).

[6] Data on commodity prices are routinely available from many sources on the web. Figures 2.5 and 2.6 are adapted from kitco.com and http://mining101.blogspot.com/2008/06/rising-copper-prices-copper-basics-bhp.html

[7] Data on changes in perceived estimates of oil reserves come from Lin (2009). See also Babajide (2009)

[8] Many descriptions of the evolution and use of UAVs are available. Much of the attention focuses on the largest versions, the Global Hawk and the Predator. Readers may find http://prometheus.med.utah.edu/~bwjones/C1311122445/E20080222191015/index.html particularly informative.

CHAPTER 3

FLEXIBILITY CAN INCREASE VALUE

"Invincibility lies in the defense; the possibility of victory in the attack."

Attributed to Sun Tzu

"Robust design is passive way to deal with uncertainty. Flexible design is active way to deal with uncertainty."

This chapter shows how flexibility in the design of projects can increase expected value, often dramatically. Flexibility in design routinely improves performance by 25% or more. It achieves such results by configuring existing technology to perform better over the range of possible future circumstances. Flexible design enables the system to avoid future downside risks and take advantage of new opportunities. By cutting losses and increasing gains over the range of possible futures, flexible design can improve overall average returns, that is, the expected value.

Moreover, flexible design often also reduces initial capital expenditures (CAPEX). It can thus obtain greater expected value at less cost, leading to substantial increases on the return on investment. Indeed, flexible designs lead to smaller and inherently less expensive initial systems. This is because the flexibility to expand as needed means that it is not necessary to build everything that might be needed at the start. Thus even though the flexible design has extra features that enable its ability to adapt to different situations, the extra cost of these features is frequently much less than the savings achieved through a smaller initial design.

The demonstration of the advantages of flexible design sets the stage for the second part of the book, the detailed presentation of the methods for identifying, choosing and implementing flexibility to maximize value to the system.

The demonstration is by example, inspired by an actual development in Southeast England. The case is transparently simple, to enable us to develop an intuitive understanding of the results. The case is however realistic. It represents in the small the same issues and opportunities as large, complex developments such as a major, $250 million project discussed at the end of the chapter.

We begin by showing that standard evaluations, using a cash flow analysis based on a most likely forecast, can give misleading, wrong results. This is because of the flaw of averages. Any analysis that does not account for the possible range of distributions of scenarios can be widely wrong – not only in its estimation of value but also, perhaps more importantly, in its identification of the best project.

The example then demonstrates that the right kind of flexibility in the design gives the project leader three kinds of advantages. It can

- Greatly increase the expected value of the project;

- Enable the system manager to control the risks, reducing the downside exposure while simultaneously increasing the upside opportunities, thus making it possible for developers to shape the risk profile. This not only gives them greater confidence in the project, but also may reduce their risk premium and thus further increase value. and
- Often significantly reduce first costs of a project – a counter-intuitive result due to the fact that the flexibility to expand easily means that many capital costs can be deferred until prospective needs can be confirmed.

Example: parking garage

The example concerns a multi-level parking garage an entrepreneur plans to build next to a new commercial center in a region whose population is growing. Details are in Box 3.1. For the purpose of the example, we assume that the only design parameter of interest is the number of levels. So the issues addressed are: What is the value of the project? How many levels is it desirable to build?

We use the example in three steps to demonstrate the value of flexibility in design.

- Base case: we look at the standard evaluation process, which evaluates a project based on the most likely forecast.
- Recognizing effects of uncertainty**:** we compare the base case with what happens when we realistically recognize that the world is uncertain. The example shows that we not only value the project quite differently from the base case, but also that the best design can turn out to be different. In short, we show that the standard process is fundamentally deficient.
- Using flexible design to manage uncertainty: we show how the results of the analysis considering uncertainty helps us understand how to improve design and value. In particular, we show how flexibility enables systems managers to adapt the project to the future as it evolves, and thereby avoid downside risks and take advantage of upside opportunities.

[Box 3.1 about here]

Base case: standard analysis using assumed projection

The standard analysis uses a straightforward discounted cash flow analysis. We set up a spreadsheet of the annual costs and revenues for any configuration of the project, sum to find the net cash flows in each period, and discount these sums using a discount rate appropriate to the enterprise. Table 3.2 shows this layout for a particular design. To find the best design – that is, the one that maximizes value – we use this analysis for all the possible levels of the garage. We find out if the project is worthwhile by seeing if the best design leads to a positive net present value (NPV). (Appendix B gives details on the procedure.)

[Tables 3.2 and 3.3 about here]

Applying the standard discounted cash flow analysis to the case, we obtain the results in Table 3.3. As it shows, the project appears to be worthwhile, delivering positive NPV beyond the required 10% cost of capital. Further, the 6-level design maximizes the NPV. In this case, as often occurs, the design has a “sweet spot” that maximizes value. The graph of NPV versus number of levels illustrates this phenomenon (Figure 3.2). If the garage is too small, it does not generate enough revenue to cover the fixed costs. If it is too large, it may not fill up sufficiently to cover the costs of the large size. The right balance between being too small and too large lies somewhere in the middle.

This analysis is standard and straightforward. The answers to the basic questions of evaluation and design appear to be that the

- Best design is to build 6 levels, and
- Garage is a worthwhile opportunity and delivers an NPV of 2.5 million.

Unfortunately, this solution is wrong on both counts. The next sections explain why these answers are incorrect.

[Figure 3.2 about here]

Recognizing effects of uncertainty

Recognizing the uncertainty: The actual evolution of demand for parking will be highly uncertain over the 15-year life of the project. Both the initial and the additional demand over the coming decade and beyond could be much lower than expected – or they could be much higher. Alternatively, there could be low initial demand and high growth or vice versa. The value of the project under any of these circumstances would not be the same as that based on the single projected estimate. In general, as indicated in the discussion of the flaw of averages, we can anticipate that the actual expected value of the project will be quite different from that calculated in the naïve case that assumes just one possible outcome.

Furthermore, as the example shows, the design that maximizes value can also be different when we look at the realistic situation that acknowledges uncertainty. Such a change in the optimal design may or may not occur. It is however a distinct possibility. As the example indicates, we cannot assume that the ranking of the alternative designs will be the same when we consider uncertainty.

To explore the effects of uncertainty in our simple example, we assume that the three parameters defining demand can be off by up to 50% either side:

- Initial demand can be anywhere between 375 and 1125 spaces;
- Demand growth to year 10 can be anywhere between 375 and 1125 spaces; and
- Demand growth after year 10 can be anywhere between 125 and 375 spaces.

As Chapter 2 documents, real outcomes can easily diverge from prior forecasts by similar amounts.

We also assume for simplicity that none of the parameters in these ranges is more likely than any other. In mathematical terms, we assume a uniform probability distribution of these parameters. The variability thus averages out to the original demand projection. This is important to avoid bias when we compare the results of the case with uncertainty with the base case using the original fixed projection. Any change in the overall valuation of the project will thus be entirely due to the variability around the projection. Using a uniform distribution ensures a fair comparison between the base case that ignores uncertainty and the realistic one that recognizes it.

These assumptions lead us to a range of different possible growth curves. Figure 3.3 shows 10 of these possibilities together with the original demand projection. Notice that the curves vary in initial demand and in demand growth. In our simplistic model, demand grows over time in all scenarios. This may not be a sensible assumption in reality. Models that are more realistic would account for the possibility that demand might dip. The assumed variability is sufficient to demonstrate the basic point: it is important to recognize uncertainty – if we do not, we get misleading valuations and rankings of projects.

[Figure 3.3 about here]

Valuation of project considering uncertainty: The valuation of a project considering uncertainty requires a simple piece of technology. This is Monte Carlo simulation. It is a computer procedure for considering the effects of the possible uncertainties fairly. In practice, the simulation

- <u>Samples</u> a large number, often 1000’s, of scenarios – for example, of demand curves from the demand distribution;
- <u>Calculates</u> corresponding realized system performances, such as the net present values of the project in our case; and
- <u>Presents</u> these results in a range of convenient forms.

Monte Carlo procedures are mathematically solid and reliable. The association between Monte Carlo and gambling and irresponsibility should not put one off. The technology is an indispensable tool for the analysis of the effects of uncertainty on design options.

Monte Carlo simulation transforms distributions of uncertain inputs into a distribution of uncertain performance by repeatedly sampling the uncertain inputs and recording the corresponding performance. Appendix E provides analytic details. Importantly, Monte Carlo simulation is efficient. A laptop computer can examine a financial spreadsheet for 1000’s of possible outcomes in just a few seconds.[1] Monte Carlo simulation thus provides the way to calculate the consequences of uncertain futures.

The evaluation of the average NPV is the immediate result of a Monte Carlo simulation. This “Expected NPV” (ENPV) is simply the average of all NPV realizations associated with each of the 1000s of simulated possibilities. This is the value we can expect to receive from the project, on average. It is the factor that directly compares with the NPV estimated using a single projection, a forecast of the average conditions. In line with the Flaw of Averages, and as Table 3.4 shows, we typically see that the ENPV differs substantially from the naïve NPV calculated around a single forecast.

[Table 3.4 and Figure 3.4 about here]

Notice in Table 3.4 that the ENPV values for the parking garage are not only different but also all lower than the naïve estimates based the single projected demand. Why is this? The explanation in this case is that higher demands for parking spaces do not increase value sufficiently to compensate for the losses in revenues corresponding to equivalent lower demands. When the parking garage is full, it cannot accommodate more cars, and cannot profit from greater demand. This capacity limitation thus systematically affects the value of the project.

Monte Carlo simulation helps us understand what is going on. The process generates the information needed to show us in detail what happens to the system. In addition to providing us with summary statistics, such as the ENPV, it generates graphs of the uncertain system performance. These illustrate what is happening. Figure 3.4 for example shows the histogram for the distribution of the net present value for a design of 6 levels, obtained from 10,000 sampled demand scenarios. It documents the fact that while the project generates a distribution of results with a downside tail of significant losses if demand is low, it does not deliver a counter-balancing upside tail of higher gains when demand is high. As Figure 3.5 shows, the wide range of possible high demands all lead to the maximum value the garage can deliver at full utilization. The distribution of high demands compacts into a narrow range of highest values, thus giving more than 30% probability of being in the highest range, as the far right of the histogram shows.[2]

It is important to recognize that this skewed distribution of performance can result from balanced distributions of inputs. This is exactly what happens in this example. The uncertainties equally distributed around the single forecast have led to the skewed results shown in Figure 3.5, and indicated in Box 3.2. What happens is that the working of the system transforms the data that comes in, often to something that is quite different in character. In this particular case the limited size of the garage, the capacity constraint, distorts the results downwards. In other cases, the distortion can be upwards. In any event, a distortion generally occurs and this is the mechanism that leads to the flaw of averages.

[Figure 3.5, Box 3.2 and Table 3.4 about here]

Notice now further that the design that maximizes value under realistic conditions is not the same as the one that appears best using a simple forecast. Indeed, the ENPV of the 6-story garage, averaged over the 10,000 demand scenario, is negative. The 5-story garage has an average NPV of 0.3M and outperforms the 6-story design. The characteristics of the system that change the distribution of the uncertainties can thereby also change the relative value of the design alternatives. As Table 3.4 highlights, when we consider uncertainty, the 5-level design now appears better – more profitable – than the 6-level design that appeared best when looking only at a single forecast!

The standard evaluation procedure thus not only leads to wrong valuations, but can also promote wrong designs. How can we identify better designs that deliver best performance under realistic conditions?

[Figure 3.7 here – set as Box 3.3]

Using flexible design to manage uncertainty

Now let us get back to the original questions:

- How can we make the most of the opportunity of a project? And
- What is the project worth?

We see that the uncertainties surrounding a project affect its performance in complicated and perhaps unexpected ways. How do we go about understanding the behavior? How do we design the project to make the most of the opportunity?

Understanding the system: A good way to understand the opportunities for improving a design is by looking at the target graph. This represents the cumulative chance of getting a result below any specific level, going from below the lowest value (which has no chance) to a result at or below the highest value (which has 100% chance). Taking any specific performance level as a target, the percentage of outcomes below this level is the chance that the design will not deliver the target value.

Figure 3.8 shows the target curve for the 6-level parking garage. It indicates, for example, that there is a 10% chance of losing about 15M or more, that is, of missing a -15M NPV target. Analysts sometimes refer to this as the 10% “Value at Risk”.[3] Reading up the curve, it also shows that this design has about a 40% chance of not breaking even, and a 60% chance that the realized NPV will be below about 5M, which is the same as a complementary 40% chance of realizing an NPV greater than 5M. Figure 3.9 shows the equivalent curve, plotted in terms of achieving a result. This risk curve is standard in the oil and gas industry and elsewhere.

[Figures 3.8, 3.9 and Box 3.4 about here]

We can develop our understanding of the opportunities for improving the expected value of a project by looking at how different designs affect the target curve. Consider Figure 3.10, which shows the target curves for designs with different levels of the parking garage. In the lower left hand side, it shows that the minimum performance increases substantially as we use smaller designs. Specifically, the probability of losing 10 million or more drops from 30% for a 7-level garage, to about 5% for a 5-level structure. The graph reveals this as a definite improvement.

More generally, ***shifting the target curves to the right improves performance.*** Moving the curve to the right means that higher targets have a higher chance of being achieved. This is a crucial paradigm for the design of systems under uncertainty. We need to find designs that achieve this improved performance.

[Figure 3.10 about here]

Now that we have found a way to articulate improve performance under uncertainty, we can try to understand why we get this result. Developing our understanding of how the system responds to uncertainties will help us develop solutions that maximize performance. In this case, the reason smaller designs are less likely to lose money is clear: if you invest less at the start, you reduce the maximum amount you can lose. Of course, this does not mean we should always invest less; spending more – for example in enabling flexibility – if often the way to increase overall expected value.

[Box 3.5 about here]

Phasing investments: We can phase investments to reduce initial cost and therefore the maximum exposure to loss. The phasing can either be fixed, in that it follows a predetermined schedule, or it can be flexible in that the later stages are optional.

Phased design can be a good idea even without the recognition of uncertainty. Deferring the cost of building capacity until it is needed saves on financing costs. At a reasonable opportunity cost of capital, 10% for example, deferring capital expense by 4 years reduces the present value of each deferred dollar to 70 cents. However, the economies of scale that designers might achieve by building capacity all at once counterbalance the potential savings of phasing to some degree.[4] The value of phasing needs careful consideration.

Based on the projected demand growth for the parking garage, it would be reasonable to plan a fixed phased expansion. The idea would be to add floors one by one, tracking the projected demand as follows:

- Begin with 4 levels;
- Expand to 5 levels in year 3;
- Expand to 6 levels in year 4, and
- Expand to 7 levels in year 7.

Note that phasing requires the original design to incorporate the ability to expand. In this case, planned phasing to 7-levels means that we would have to build the first 4-level phase with extra large columns and foundations, sufficient to carry the eventual loads of a higher structure. We assume this cost adds 30% of the cost of the first two levels.

From the perspective of a standard NPV analysis based on projected demand, this fixed phasing more than doubles the NPV of the project, from 2.5M for a fixed 6-level design to 5.4M for the phased design (see Table 3.5). From the naïve perspective therefore, a fixed phasing of the development of the parking garage seems highly attractive. But how does it compare when we value it recognizing uncertainty?

[Table 3.5 about here]

The expected NPV for the fixed phased design when we recognize uncertainty is greatly different from the value based on the single projection. In fact, when we test the design with 10,000 demand scenarios, the realized ENPV averages at -0.4M (see Table 3.5). This realistic assessment shows that this fixed phased design is actually worse than the best fixed designs of 5 or 6 levels – contrary to what the standard analysis indicated.

To understand what is happening, we can look at the target curve comparing the fixed phased design with the better 5-level fixed alternative (Figure 3.11). The fixed phased design curve is far to the left for lower values. This means that it is much more likely to lead to major losses. This is because the fixed phased design locks into an eventually larger garage – up to 7 levels. If demand increases as projected, this plan leads to good results. Hence, the target curve for the phased design curve is further to the right than the corresponding part of the fixed design curve. However, if demand is less than expected, the commitment to build large is obviously wasteful. Locking the system into an expansion plan is not a good idea when the need does not materialize as initially expected.

[Figure 3.11 about here]

We can also use the target curves to identify the probability that any particular design will "break even", that is, repay the present value of the investment costs. This break-even probability is the complement of the probability that the project has an NPV of zero or less. Thus Table 3.5 shows the break-even probability for the phased design to be 65 %, which is 1.0 less the intercept of its target curve with NPV=0 in Figure 3.11.

Flexibility is the way: The problem with the fixed phased design is that it locks into an expansion plan that may not make sense. This observation points to the preferred solution. A flexible design permits but does not require expansion. A flexible design thus positions the system so that it can expand when it makes sense to do so, but does not commit the managers to expand if the actual situation that occurs does not justify greater capacity.

A flexible design recognizes that we will learn more about system requirements during the lifetime of the project. Once we open the system, we will quickly know the initial demand. As we gain experience, we will also be able to get a better idea of the real potential demand growth. We may consequently postpone or accelerate the expansion; we may expand less or not at all, or may expand more.

Proper development of a flexible design requires reasonable criteria for expansion. Depending on the situation, managers may want to expand somewhat in anticipation of demand, to avoid any lost sales. Alternatively, they may wish to defer expansion until a pattern of higher demands is well established. Judgment about which criteria to apply is a management decision. From an analytic perspective, we can easily program any management decision rule into the analysis.[5] We can therefore anticipate the range of outcomes and the future performance of the system.

The flexible design can indeed be far superior. Consider for example the flexible design of the garage with an initial build of 4 levels that incorporates the ability to expand incrementally up to 9 levels. To value this design, we have to agree on a criterion for expansions, a contingency plan. In this example, our decision rule was to add another level whenever demand exceeds capacity for two years in a row. As Table 3.6 shows, the flexible design is preferable on all counts to the fixed alternatives.

[Table 3.6 about here]

The flexible design delivers much greater Expected Net Present Value than the best alternative fixed design, 2.1M compared to 0.3M. This kind of result is common, as case examples demonstrate throughout the rest of the book. By building small, flexible design lowers the risk of bad performance: less invested means less can be lost. By incorporating the ability to expand easily, it allows the system managers to take advantage of favorable opportunities whenever they appear. Both actions lead to greater expected value. Correspondingly, they also lower the Value at Risk and raise the Value at Gain. The target curves in Figure 3.12 shows this effect graphically. The NPV target curve for the 4-level flexible design mirrors the performance of the 5-level fixed design in low demand scenarios but moves to the curve for the 6-level fixed design in high demand scenarios. Both avoiding losses and increasing possible gains move the target curve the right, towards better performance.

[Figure 3.12 and Box 3.6 about here]

Most remarkably, flexible designs often **cost less than the inflexible designs.** This result is truly important that runs against the intuition that flexibility always costs more. To appreciate this fact, it is important to create a fair comparison between the designs that we might implement: the flexible or the inflexible design. The basis for comparison is the initial investment we might make, between the initial capital investments (CAPEX). The comparison needs to be between what one might actually decide, not what the structure looks like. Thus, we need to compare the following possible decisions:

- The choice to build an inflexible design means that we have to build it larger than immediately needed, so that we can benefit from future growth -- thus the 5-level solution;
- The choice of the flexible design allows us to build only for immediate needs, as this design allows us to add levels as needed – thus the 4-level solution.

In short, the fair comparison is between the alternative investments, not between two 4-level or two 5-level designs. The fair comparison is between the 4-level flexible and the 5-level inflexible design.

For this case, the initial cost of the flexible design is 10% less expensive than the competitive 5-level design. This is because the flexible designs allow the system managers to build small initially. This can create tremendous savings, far greater than the cost of enabling the flexibility. In this case, the extra cost of building stronger columns and footing to enable expansion is less than the savings resulting from building the structure smaller initially. Flexibility can be a "win-win" design!

Add a couple of target curves for some realistic cases. Yang (automobile manufacturing), perhaps Dykes (wind energy). In any case, want to show increases in value!

Application to a major project: Health Care Service Corporation building

The parking garage case is a simple example, but a realistic case that scales up for major projects. The flexible design of the Health Care Service Corporation (HCSC) building in Chicago is one of several recent examples.[6]

The HCSC building is certainy a major project. As Wittels and Pearson (2008) reported,

> "It has a footprint of about 36,000 square feet and the original 30-story building had a gross floor area of 1,4 million square feet. It has the second largest footprint of an office building in Chicago, after the Willis (formerly Sears) Tower, which is currently the tallest building in the United States, and was for many years tallest in the world."

The pictures in Figure 3.13 document the comparison. The Willis Tower is to the left.

The Health Care Service Corportation designed its headquarters for flexible expansion. The flexible design called for a first phase of 30 stories with the ability to expand to 54 stories if and when this proved to be desirable. HCSC anticipated that it would grow nationally, in line with national trends powered by economies of scale in the managed care industry in the United States. When HCSC planned and committed to the original design in the early 1990s, it could of course not be sure what its situation would be 15 years or more in the future. Thus it built for immediate needs, with the capability to expand if and when this would be needed.

The flexible design had to be carefully organized for vertical expansion. Most obviously, the initial structure was designed to carry about twice the weight. It also needed the strength to resist much higher sideways bending forces from the greater wind pressures on a taller column. Also, all the services and conduits had to be designed for twice as many people. The original design had to provide double the number of elevator shafts, for example. Finally, the architects had to define how they would organize the construction of the 24-story building when they had to pass all materials and work crews to the top of a fully occupied building.

HSCS saved money by building its headquarters with the flexibility to expand vertically. It did so because it could confidently build a smaller building from the start, knowing that it could accommodate growth smoothly. If HCSC had not had this expansion capability, it would have needed to build a much higher building in the 1990s. Viewed as a 30-story building, the flexible design cost more than than an equivalent structure, but this is not a fair comparison. The flexible 30-story design needs to be compared to the larger, perhaps 40-story building that was the realistic alternative at the time of decision in the 1990s. Compared to the actual choices, the flexible 30-story building was less expensive.

The construction of the second phase, involving 24 stories, began in 2007 for scheduled completion in 2010. The work was done while the building was continuously occupied by a work force of about 3000 people.

Take-away

Flexibility is the way. It is an effective way to improve the expected performance of systems in uncertain environments. The gains to be made can be impressive. This is especially so when the flexible design also reduces the initial capital expenditure required for the project.

The type of flexibilities that are most useful in achieving this will of course depend on the context. Project or system designers face the challenge of bringing their technological expertise to bear on the creation of appropriately suitalbe systems that can cope with many unfolding futures.

The correct way to value projects requires the recognition of uncertainty. Evaluations done around average or most likely forecasts of future needs or requirements will systematically lead to incorrect answers, both as regards the level of value and the ranking of the choices for design. Proper analyses require an understanding of the effects of uncertainty.

Target curves representing the uncertain performance of a system provide an effective, systematic tool to gain insight into the behavior of alternate designs. They highlight the risks and opportunities of different designs. They give meaning to the notion of “improving” the uncertain performance of a system: The goal is to use creative design to move the performance curve, or at least substantial parts of it, to the right, to reduce downside risks and increase the opportunites for benefiting from the upside.

Box 3.1

Parking garage details[7]

The developer can lease the land for 15 years, after which the land and the garage revert to the landowner. The cost would be $3,330,000 per year.

The site can accommodate 250 cars per level. Due to variation in traffic, we assume that the average utilization of capacity over the year does not exceed 80% of installed capacity, i.e., the effective capacity per level is 200 cars. The construction consultant estimates that given the local conditions it will cost $17,000 per space for pre-cast construction, with a 10% increase for every level above 2, that is, beyond the ground level and the 1st level up. This premium covers the extra strength in the columns and foundations to carry the load of additional levels of the structure and the cars.

The average annual revenue per space used will be $10,000, and the average annual operating costs (staff, cleaning, etc.) about $3,000 for each space. The entrepreneur uses a 10% discount rate.

An expert consultant projected demand for parking spaces over the next 15 years. Based upon local conditions, future demand is assumed to be determined by 3 parameters:

- Initial demand for parking during the first year of operation (750 spaces);
- Demand growth during the first 10 years of operation (an additional 750 spaces, leading to a demand of 1500 spaces in year 10); and
- Demand growth after year 10 (an additional 250 spaces, leading to a saturation demand of 1750 spaces).

These demand parameters would typically be estimated using demographic projections and past data from similar projects, and then used to define a ramp-up curve to the final demand of 1750 spaces. Our example assumes this curve is of the form

$$d(t) = 1{,}750 - a \cdot \exp(-bt)$$

with a and b chosen to match the year 1 and year 10 demands. This gives the demand curve over time shown in Figure 3.1 and Table 3.1.

[Figure 3.1 and Table 3.1 here]

Box 3.2

System changes the distribution of uncertainties

In the case of the parking garage, the capacity of the facility limits the maximum use and therefore the upside potential of the facility.

[Figure 3.6 here]

Box 3.3

Figure 3.7 here

Box 3.4

Reverse percentile graph

The reverse percentile graph is an alternate to the target curve. If displays the chance of achieving a certain level of performance, rather than missing it. Thus, it shows a 100% chance of getting the lowest level, and ultimately no chance of exceeding the maximum level.

Some people find the reverse percentile graph more intuitive. In some industries, it is the standard way to present data. **[Check what this curve/graph is called in the oil industry]**

The information in the reverse percentile graph for the parking garage, shown below, is the same as in the target curve. In this case, there is about a 40% chance of realising an NPV of 4M or more, a 60% chance of breaking even, and a 90% chance that a loss does not exceed 15M.

Box 3.5

KEY PARADIGM

IMPROVING A SYSTEM IN AN UNCERTAIN ENVIRONMENT MEANS
MOVING ITS TARGET CURVE TO THE RIGHT,
THUS MAKING THE ACHIEVEMENT OF TARGETS MORE LIKELY.

Box 3.6

Summary of calculations for parking garage example

[Table 3.7 here]

Table 3.1: Projected growth in demand for parking

Year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Demand (spaces)	750	893	1015	1120	1210	1287	1353	1410	1458	1500	1536	1566	1593	1615	1634

Table 3.2: Spreadsheet for garage with 6 levels, based on projected demand (costs and revenues in millions)

Year	**0**	**1**	**2**	**…**	**15**
Demand (spaces)	0	750	893	…	1,634
Capacity (spaces)	0	1,200	1,200	…	1,200
Revenue	0.0	7.5	8.9	…	12.0
Operating costs	0.0	3.6	3.6	…	3.6
Land leasing and fixed costs	3.3	3.3	3.3	…	3.3
Cash flow – actual	-3.3	0.6	2.0	…	5.1
Cash flow – discounted	-3.3	0.5	1.7	…	1.2
Cash flow – present value	26.7				
Capacity cost < two levels	6.8				
Capacity cost > two levels	17.4				
Net present value	**2.5**				

Table 3.3: Value in millions of different designs of the parking garage, based on projected demand

Design Levels	**NPV**
4	- 1.2
5	2.2
6	**2.5**
7	- 0.7

Table 3.4: Comparison of actual expected net present value in millions of project (ENPV) and value estimated using single projected demand

Design Levels	NPV	Expected NPV
	Based on demand	
	Projected	Uncertain
4	- 1.2	- 2.5
5	2.2	0.3
6	2.5	- 0.1
7	- 0.7	- 3.8

Table 3.5: Evaluation of phased demand, values in millions

Design Levels	NPV	Expected NPV	Break-even probability
	Based on demand		
	Projected	Uncertain	percent
4	- 1.2	- 2.5	0
5	2.2	**0.3**	64
6	**2.5**	- 0.1	59
7	- 0.7	- 3.8	48
Phased	5.4	- 0.4	66

Table 3.6: Comparison of flexible design with alternatives, values in millions

Design Levels	Initial capital outlay	NPV	Expected NPV	Break-even probability	10% value	
		Based on demand			at Risk	to Gain
		Projected	Uncertain	percent		
5	19.2	2.2	0.3	64	-7.6	5.4
6	24.2	2.5	- 0.1	59	-14.3	9.6
4, flexible expansion	16.7	n/a	2.1	66	- 7.4	9.2
Which best?	**Flexible**	**n/a**	**Flexible**	**Flexible**	**Flexible**	**Flexible**

Table 3.7: Summary of results for parking garage example, values in millions

Design Levels	Initial capital outlay	NPV	Expected NPV	Break-even probability	10% value	
		Based on demand			at Risk	to Gain
		Projected	Uncertain	Percent		
4	14.7	- 1.2	- 2.5	0	- 6.3	- 0.7
5	19.2	2.2	0.3	65	- 7.6	5.4
6	24.2	2.5	- 0.1	60	- 14.3	9.6
7	29.6	- 0.7	- 3.8	45	- 23.5	10.7
Phased, planned	16.7	5.4	- 0.4	65	- 16.2	6.6
3, flexible	12.6	n/a	0.8	60	- 5.9	5.2
4, flexible	16.7	n/a	2.1	65	- 7.4	9.2
5, flexible	21.2	n/a	1.6	60	- 10.8	11.5

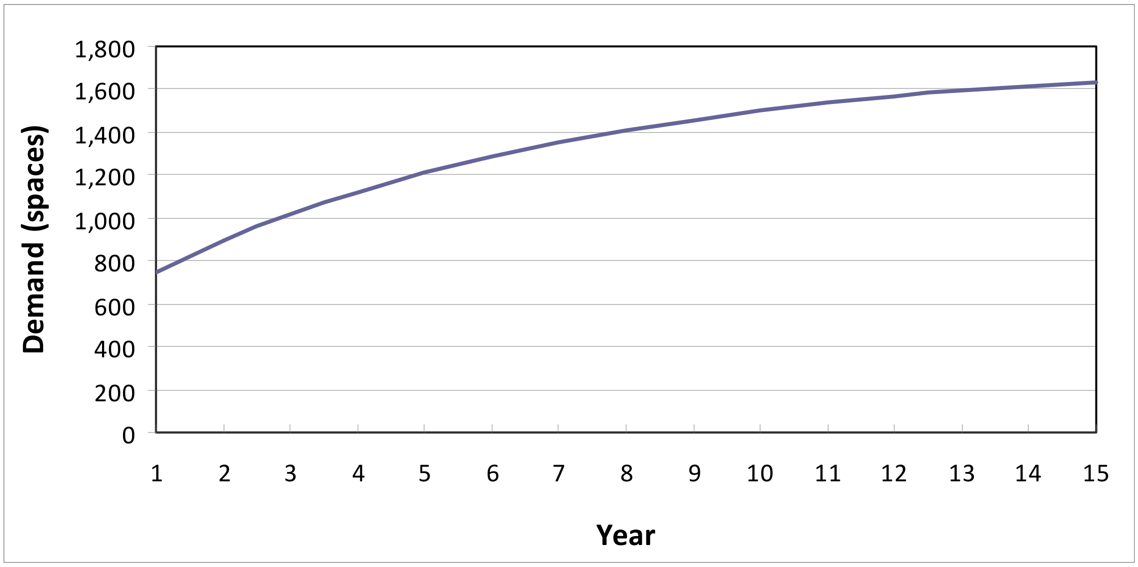


Figure 3.1: Projected growth in demand for parking

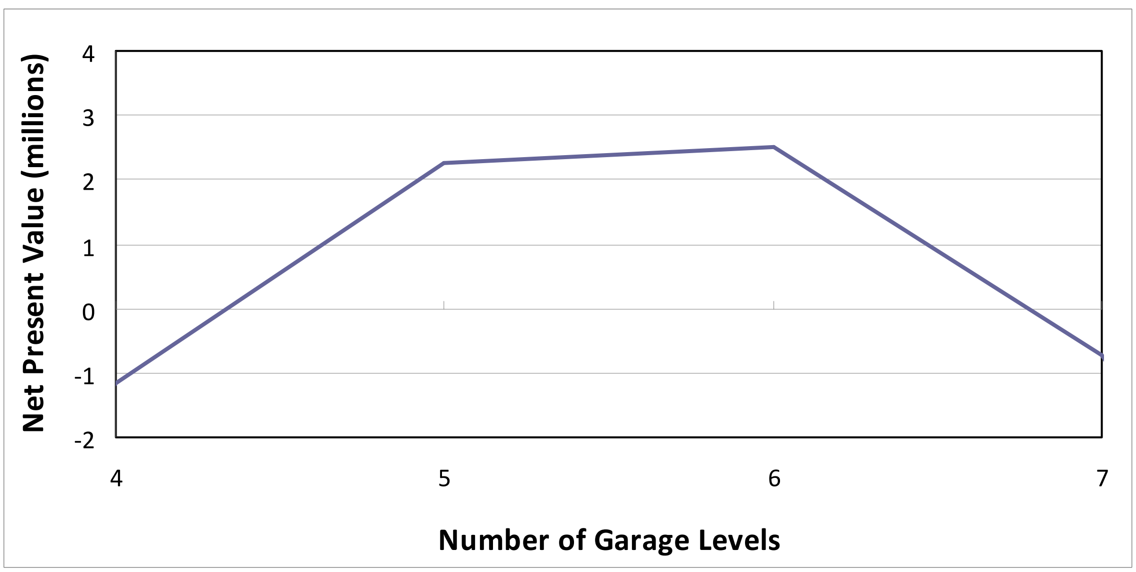


Figure 3.2: The NPV associated with different designs, assuming deterministic demand.

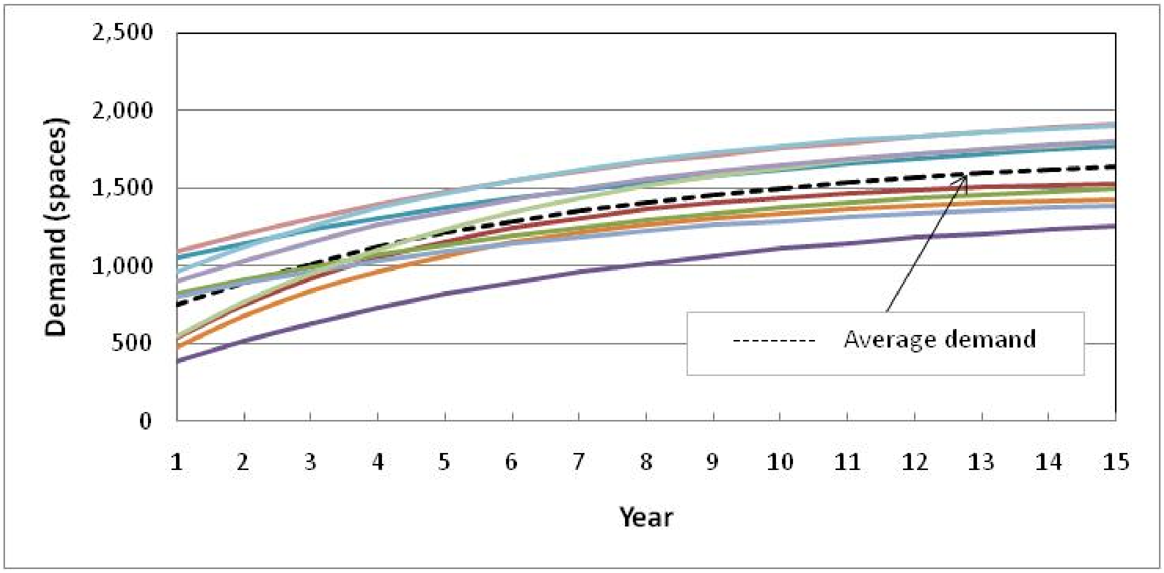


Figure 3.3: Some possible curves of demand growth, compared to original single projection.

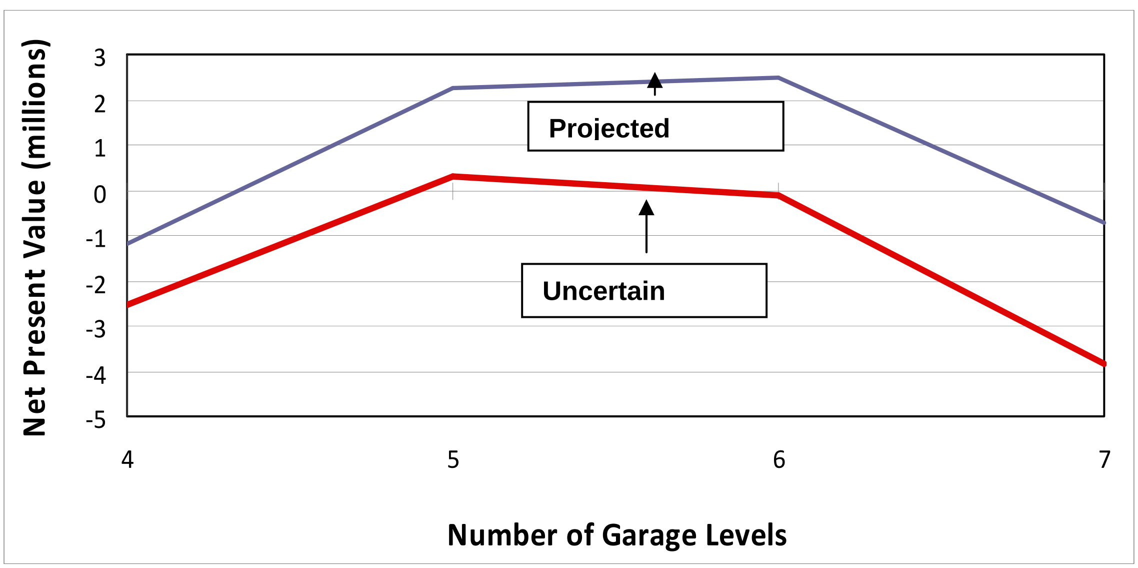


Figure 3.4: Comparison of the actual Expected NPV associated with the uncertain demand and the NPV calculated for a deterministic demand, for different designs.

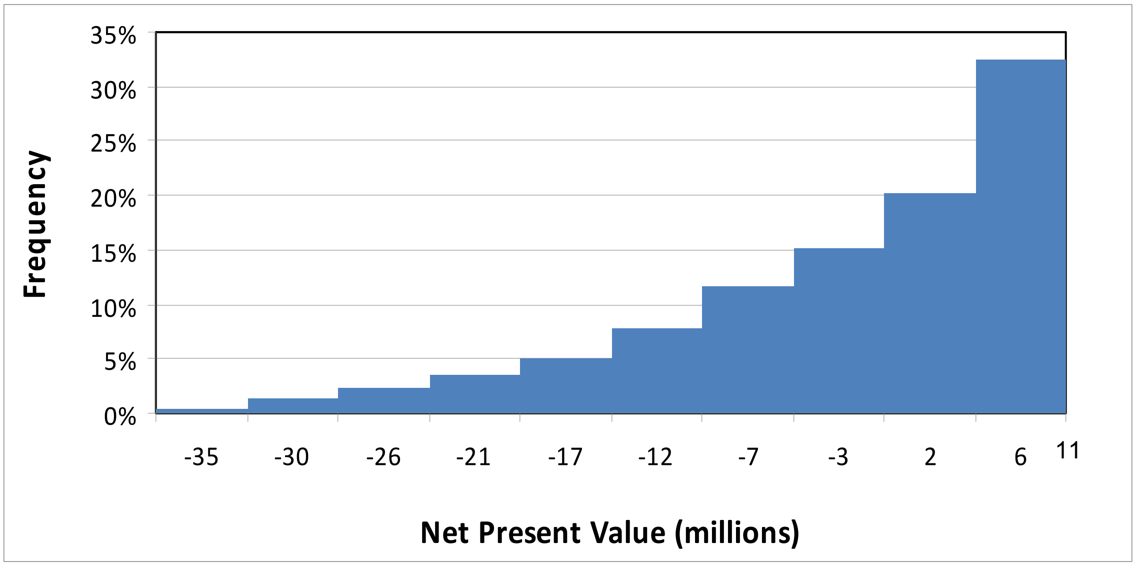
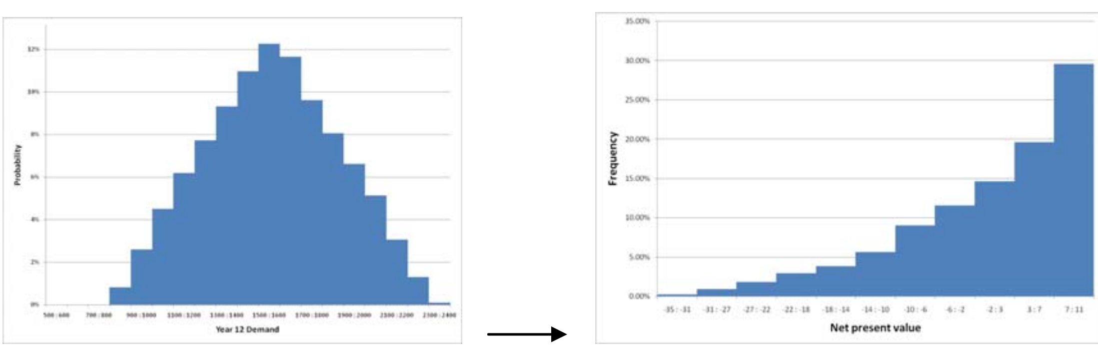


Figure 3.5: Histogram showing the distribution of possible values of the fixed, 5-story parking garage, considering the possible uncertainty in demand.

Note – the numbers on the horizontal axis should delimit the ends of the ranges. We need to fix this.



Symmetrical Input Distribution

Skewed Distribution of System Performance

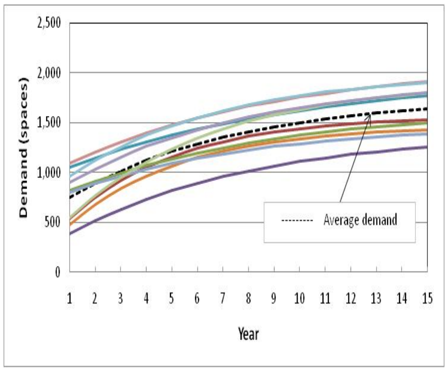
Figure 3.6: -- to be set in Box 3.2, no extra caption

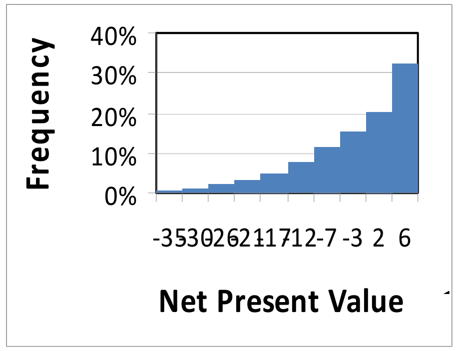
KEY TECHNOLOGY

MONTE CARLO SIMULATION PROCESSES

DISTRIBUTIONS OF UNCERTAIN INPUTS AND PRODUCES

DISTRIBUTIONS OF UNCERTAIN PERFORMANCE





Uncertain Environment → **Uncertain System Performance**

Figure 3.7 -- to be set as Box, no extra caption

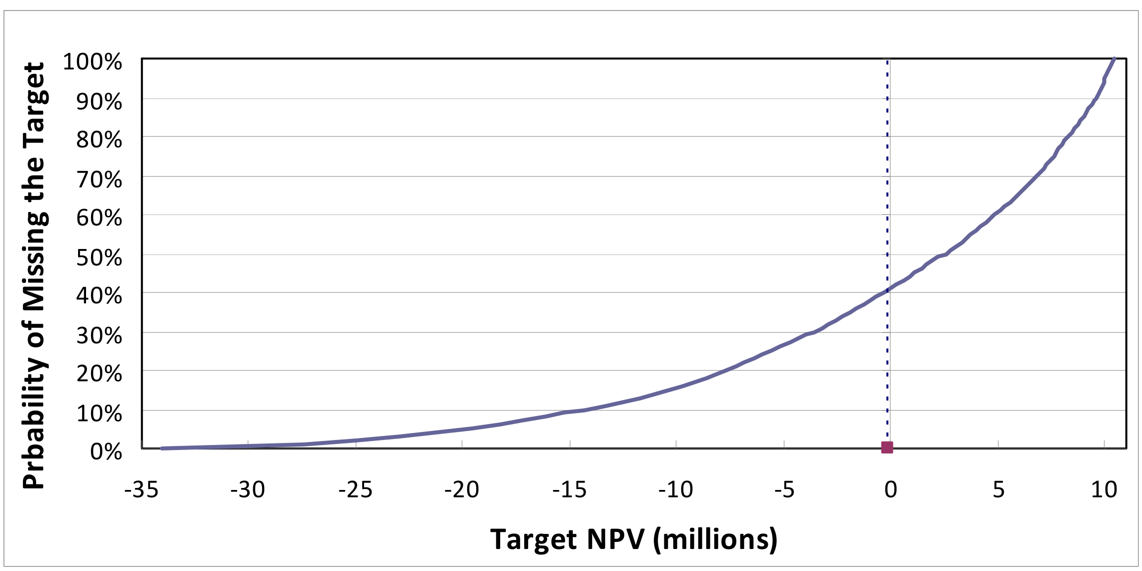


Figure 3.8: Target or cumulative percentile curve, indicating the probability of missing any target on the horizontal axis

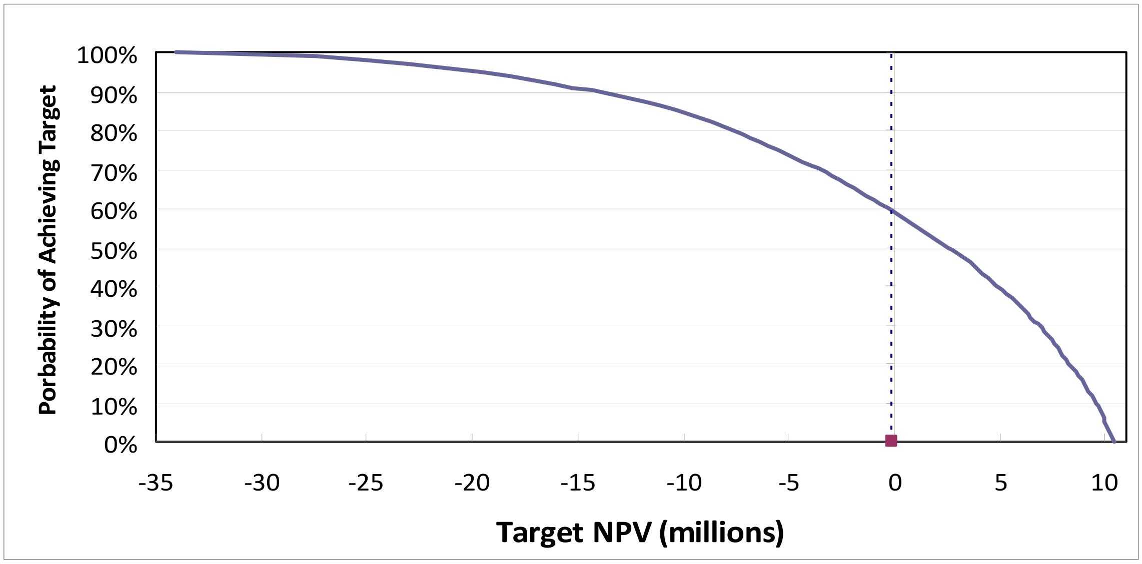


Figure 3.9: Risk curve or reverse percentile curve showing the probability that the system will meet or exceed the performance indicated on the horizontal axis

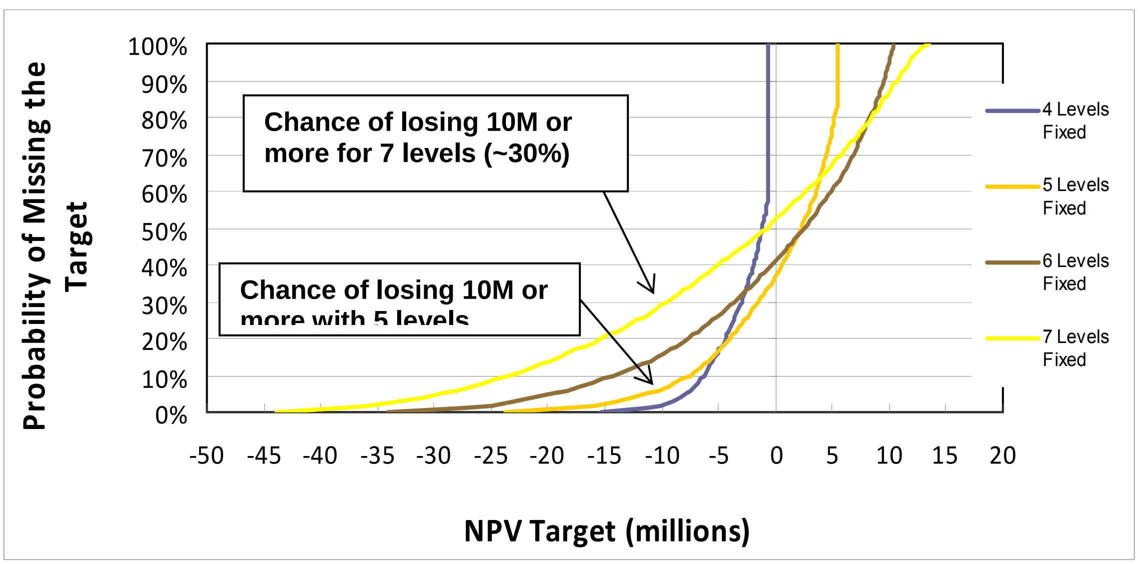


Figure 3.10: Target curves for designs with different levels of the parking garage

Chance of loosing $10M or more for 7

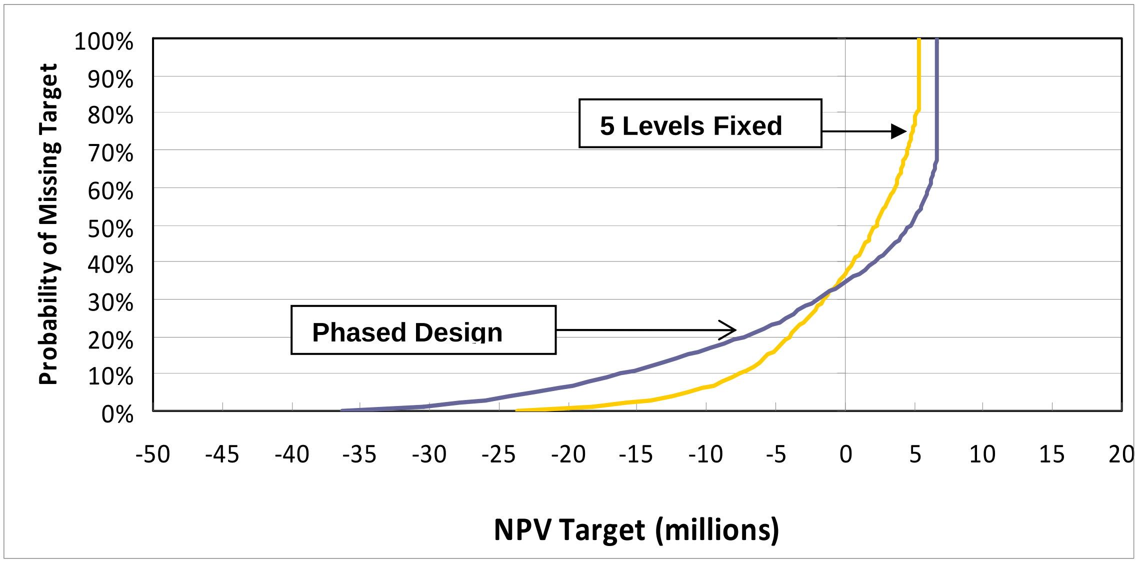


Figure 3.11: Target curve comparing phased design with fixed design with 5 levels

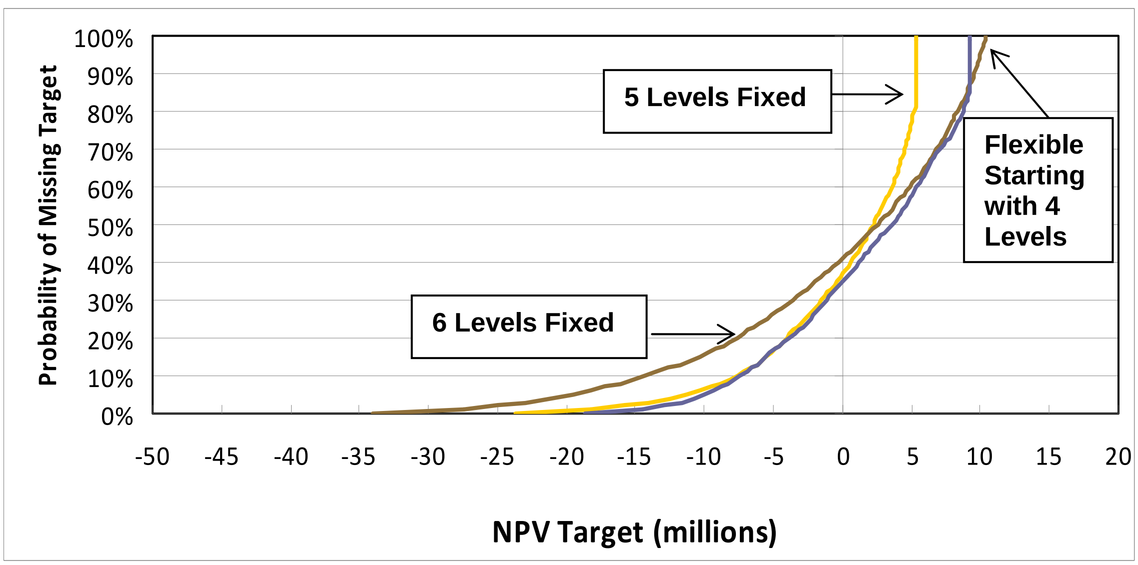


Figure 3.12: Target curve for flexible design compared to the best fixed alternatives

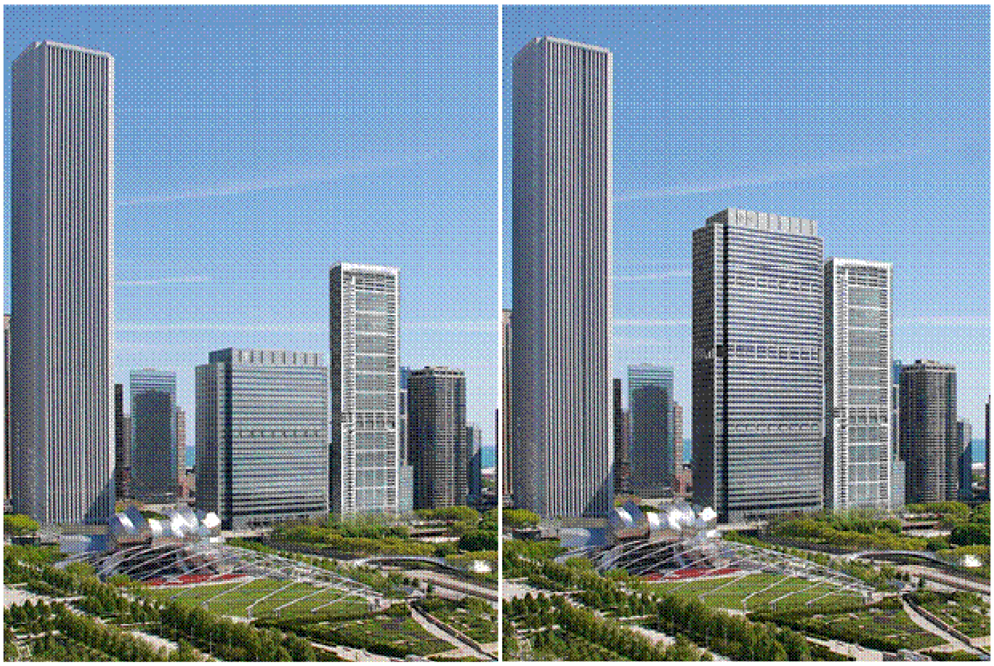


Figure 3.13: Vertical expansion of Health Care Service Corporation building in Chicago in center of image. Phase 1 (left) and Phase 2 (right)
(Source: Goettsch Partners release to Wittels and Pearson, 2008).

[1] Several open source and commercial software packages are available to perform these calculations. We used XLSIM to calculate and graph the results in this book.

[2] A histogram divides the range of a distribution into subdivisions of equal size, counts the number of occurrences of an outcome within each smaller range, and shows this number as a vertical bar. This process shows the frequency of ranges of outcomes, which is the probability distribution. In the example, part of the Monte Carlo software took the range of the 10,000 realized NPVs (approximately between -35M and +11M) and divided it into 10 regions of equal width, about 4.6M in this case. The height of the bar associated with each bin measures the fraction of the 10,000 realized NPVs that fall into the respective range. For example, about 5% of the sampled NPVs fell between -17M and -12M; about 32% between 6M and 11M.

[3] The concept of "Value at Risk" or VAR has been widely used in finance. It refers to the amount that might be lost, or the target that the design might not attain, with a specified probability. The Value at Gain" or VAG is the complementary probability of what might be gained. The cumulative probability curve is thus sometimes referred to as the VARG or Value-at-Risk-and-Gain" curve.

[4] The economies that designers can achieve by building capacity all at once include both economies of scale and the costs avoided by not working on expanding a facility that is in operation. Unless designers carefully plan future expansion, it can be problematic to construct around on-going operations. As regards economies of scale, Manne (1967) clearly presents the interplay between the discount rate (a higher one motivates deferring development), the degree of economies of scale (greater savings by building large encourage early development) and the rate of growth, which determines the size of expansion increments.

[5] See Cardin (2007) for a discussion of the effect of many different decision rules for expansion.

[6] For details see Guma et al (2009), Guma (2008) and Wittels and Pearson (2008).

[7] The case is extrapolated from an actual project in the Bluewater development in England (http://www.bluewater.co.uk/). The numbers used are representative, but do not correspond to those of the actual project. Chapters 1 and 6 also mention this garage case. A version of it appears in de Neufville and Scholtes (2006).