Australian Government

Department of Transport and Regional Services

Bureau of Transport and Regional Economics

Report 110

Risk in cost–benefit analysis

Report 110

Risk in cost–benefit analysis

ISSN 1440-9585

ISBN 1-877081-78-7

This publication is available free of charge from the Bureau of Transport and Regional Economics, GPO Box 501, Canberra, ACT 2601, by downloading it from our website (see below), by phone (02) 6274 7210, fax (02) 6274 6816 or email: *btre@dotars.gov.au*

http://www.btre.gov.au

Disclaimers

The BTRE seeks to publish its work to the highest professional standards.

However, it cannot accept responsibility for any consequences arising from the use of information herein. Readers should rely on their own skill and judgement in applying any information or analysis to particular issues or circumstances.

Printed by the Department of Transport and Regional Services

Desktop publishing by Jodi Hood

Foreword

Facts and furphies in benefit–cost analysis: transport (BTE 1999) addresses a wide range of common misconceptions about cost–benefit analysis (CBA) and its application. One is that the discount rate should include a risk premium. The report showed that raising the discount rate to account for risk could distort ranking of projects. Where projects are being ranked in order to allocate program funding and a single discount rate is being used, it would be more appropriate to discount at the long-term government bond rate, that is, the risk-free rate.

Such a conclusion may seem surprising in view of private-sector practices that insist upon higher returns from more risky investments. BTRE saw a need for the reasons behind this seeming inconsistency between good practice in the private and public sectors to be explored further, if the arguments for using a risk-free discount rate in CBA are to be persuasive.

Against this background, BTRE commissioned Professor John Quiggin of the Faculty of Economics and Commerce, Australian National University (and now at the University of Queensland) to write the paper that is reproduced as an appendix to this report. The paper was reviewed by Professor Peter Forsyth of the Department of Economics at Monash University.

Dr Mark Harvey has prepared this report in order to present Professor Quiggin's arguments to a wider audience and to spell out what they mean for practical applications of CBA. In a number of areas, the issues have been taken further, such as estimation of expected benefit–cost ratios and internal rates of return, when and how to estimate certainty equivalents, and implications for risk management. Phil Potterton supervised the project.

Professor Quiggin and Dr Harvey presented a seminar in Canberra in April 2004 to an audience that included economists employed in the government, consulting and academic sectors. The report has benefited from comments received at the seminar and in response to the draft report circulated.

Phil Potterton
Executive Director
April 2005

Contents

At a glance...

- Adding a risk premium to the discount rate to adjust for risk in cost–benefit analyses (CBAs) of public sector projects can distort project rankings. It alters costs and benefits according to a particular pattern over time, which will be correct only under assumptions that would rarely hold in practice.
- 'Downside risk', that is, biased estimates of costs and benefits due to failure to consider what can go wrong, is best addressed by using the 'state-contingent' approach, that is, assigning probabilities to different possible outcomes and estimating expected values.
- 'Pure risk', the risk that remains after removal of downside risk, can be ignored in most cases.
- Pure risk due to random variation (idiosyncratic risk), should be largely diversified away, as long as the benefits of individual projects are spread widely over large numbers of individuals and there are numerous projects.
- There remains the pure risk not diversified away because it arises from correlation between project benefits and general economic activity (systematic risk). For the typical public-sector project, the required adjustment to benefits turns out to be very small in relation to the margin for error in CBAs.
- However, where a single project has a large impact on the welfare of a small number of individuals, explicit adjustments should be made to benefits, making assumptions about the risk averseness of beneficiaries.
- Evaluation of public-sector projects at a risk-free discount rate significantly lower than rates used by the private sector for financial analysis could raise concerns about government investment crowding out private-sector investment. However, addressing downside risk for public-sector projects should work in the opposite direction. Also, the private sector has other offsetting advantages; and overall levels of government investment are budget-constrained.
- Downside risk is part of a wider problem called 'optimism bias'. Comprehensive and transparent risk assessment in cost–benefit analysis as advocated in this report should do much to counter optimism bias.

Table of Abbreviations

AADT	Average Annual Daily Traffic
ADB	Asia Development Bank
ATC	Australian Transport Council
BCR	Benefit–Cost Ratio
BDOT	British Department of Transport
BTE	Bureau of Transport Economics
BTRE	Bureau of Transport and Regional Economics
CAPM	Capital Asset Pricing Model
CBA	Cost–Benefit Analysis
CCAPM	Consumption Capital Asset Pricing Model
FYRR	First Year Rate of Return
GDP	Gross Domestic Product
IRR	Internal Rate of Return
NPV	Net Present Value
SOC	Social Opportunity Cost
STP	Social Time Preference

Executive Summary

For cost–benefit analyses (CBA) of public-sector projects, a common misconception is that the discount rate should include a risk premium in consonance with the private-sector practice of doing so.

In examining the issue, this report addresses different types of risk separately:

- downside risk, which arises from optimistic bias in forecasts; and
- pure risk, which is the variation remaining around the mean after removing downside biases. Pure risk is divided into two further sub-categories:
  - idiosyncratic risk, which is random variation; and
  - systematic risk, which is variation correlated with the level of general economic activity.

Adding a risk premium to the discount rate is a very poor way to correct for *downside risk*. It engenders little or no increase in construction costs and reduces benefits at an increasing rate with time. It would be pure coincidence if the pattern of reductions in benefits arising from a risk premium corresponded with the adjustments necessary to remove downside risk.

As long as the benefits of individual projects are spread widely over large numbers of individuals and there are numerous projects, *idiosyncratic risk* should be largely diversified away.

Adding a risk premium to the discount rate can adjust for *systematic risk* under assumptions that hold approximately for a private investor purchasing shares. However, the necessary assumptions are unlikely to hold for public-sector projects.

Risk premiums can be estimated for systematic risk as direct adjustments to project costs and benefits. However, for the typical public-sector project, the premium turns out to be very small in relation to the margin for error in CBAs. The reason is that neither aggregate consumption nor the benefits of the typical public-sector project are subject to a great deal of variability over time. Hence, for most projects in practice, pure risk can reasonably be ignored. The exceptional case is where a single project has potentially large impacts on the welfare of a small number of individuals. Even in this case, adding a risk premium to the discount rate is not the answer.

Other arguments for high discount rates include matching the social opportunity cost of capital and adjusting for the economic efficiency losses of increased taxation to finance investments. The social cost of capital is the pre-tax rate of return earned by marginal private sector investment. It would only be the correct discount rate to use when all funds for a public sector project came at the expense of private sector investment and all benefits were reinvested in the private sector. Other sources of funds are deferred consumption and overseas borrowing, which are associated with lower values for the discount rate.

Just as raising the discount rate is a poor way to adjust for risk, it is also a poor way to adjust for the economic efficiency losses of taxation. Government investment decisions generally take place in a budget-constrained environment. Hence many economically warranted projects are not implemented. Government budgetary processes, it could be argued, provide an arena in which the cost of higher taxation can be weighed against the benefits of greater public-sector investment.

Given that pure risk can be ignored in most practical situations, there remains a need to minimise downside risk. The way to do so is via the 'state-contingent' approach. It involves identifying alternative 'states of nature' in which levels of costs and benefits may be different, assigning probabilities to those states of nature, and estimating expected values for the various CBA results. Use of the state-contingent approach disciplines the analyst to consider in detail what can go wrong and to assess the impacts for the CBA.

For the rare situations where a single project has a large potential impact on the welfare of a small number of individuals, benefits accruing to those individuals need to be converted to 'certainty equivalents' — the certain monetary value that is equivalent in value to the risky benefit.

The report's conclusion that pure risk can reasonably be ignored in most situations has implications for risk management. Alternative risk management strategies can be compared using the state-contingent approach to find the one that yields the highest expected NPV. This greatly simplifies comparisons between project options having different levels of risks and costs. There is no need to estimate certainty equivalents with the requirement to make subjective judgements about the degree of risk aversity.

One strategy for managing risk is project deferment. Where a project's costs or benefits are contingent upon some future uncertain one-off event, the 'wait-and-see' option can be tested to determine whether project deferral yields a higher expected NPV.

Just as the discount rate used in CBAs represents the cost of capital to society as a whole, the discount rate used in financial analysis of private-sector projects represents the cost of capital to a single entity. The discount rate used for financial analysis will include an 'equity premium' on equity capital, and a risk premium on borrowed funds to compensate lenders for the risk of loss of capital.

Evaluation of public-sector projects at a risk-free discount rate that is significantly lower than rates used by the private sector for financial analysis could raise concerns about government investment crowding out private-sector investment. First, addressing downside risk for public-sector projects should have an offsetting effect to the lower discount rate. Second, the risks and costs for the same project are likely to be different depending on whether the project is undertaken by the public or private sectors. The two sectors have different relative advantages and disadvantages and the cost of capital is by no means the only factor that determines the net worth of a project. Third, overall levels of government investment are regulated by budgetary and political processes, not just the level of the discount rate.

Downside risk is part of a wider problem called 'optimism bias'. Besides the simple failure to consider what can go wrong, there are political–institutional factors that give project proponents incentives to overstate the positives and understate the negatives. Comprehensive and transparent risk assessment in cost–benefit analysis as advocated in this report should do much to counter optimism bias. However, it will be more effective if introduced in combination with other strategies aimed at addressing the problem.

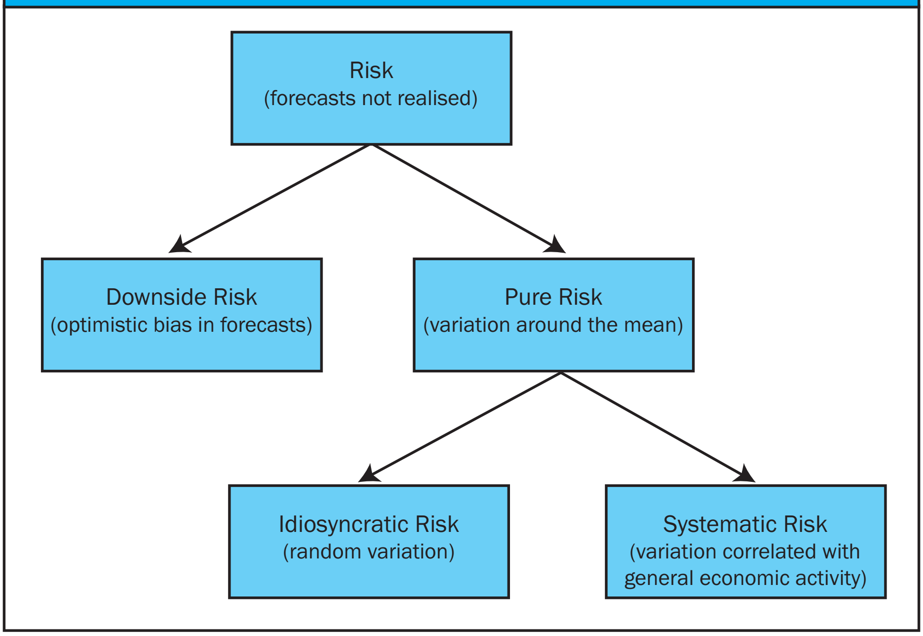
Introduction

Risk and uncertainty are cause for concern for private and public-sector investors alike. Private-sector investors insist on higher levels of expected returns to compensate them for bearing higher levels of risk. They do this by adding a risk premium onto the discount rate. Transfer of this approach to risk across to the public sector can lead to distorted results when comparing investment projects. If a risk premium is not suitable, how then should the public sector deal with risk in project evaluation?

To address this question, different types of risk are identified and considered separately. The taxonomy employed is shown in figure 1.

The suitability of the risk premium approach for cost–benefit analyses (CBAs) of public-sector projects is discussed for each type of risk, and correct approaches to dealing with risk are put forward. Numerical examples are employed to illustrate the approaches described.

FIGURE 1 RISK TAXONOMY



It is argued that the adding a risk premium to the discount rate in CBAs is a correct approach only under restrictive assumptions that would rarely hold in practice. Downside risk can be mitigated by following the 'state-contingent approach'. The approach entails identifying circumstances and amounts by which costs and benefits might differ from forecast values and attaching probabilities to the different possible values. Expected values can then be estimated for the net present value (NPV), benefit–cost ratio (BCR) and internal rate of return (IRR). Pure risk can be ignored for practical purposes most of the time. The exception is where the project may have a significant impact on the welfare of a small group of individuals. In this exceptional case, the recommended approach is to estimate a certainty equivalent. The way to do so is described below.

The state-contingent approach has applications to risk management. It can be employed to compare project options having different levels of risks and net benefits. Some options may involve reducing risk through deferment of the project until such time as future uncertain events affecting the project have unfolded. Use of the state-contingent methodology to evaluate project deferral is illustrated with a simple numerical example.

The reports also looks at some of the other reasons put forward for using high discount rates in addition to the need for a risk premium.

Discounting public-sector projects at a risk-free rate could be seen as advantaging public-sector investment relative to private-sector investment. The report presents arguments as to why this is not necessarily the case.

The final section lists some other strategies to use in conjunction with the state-contingent approach to reduce 'optimism bias' in project appraisal.

Nature of risk

All costs and benefits that go into a CBA are forecasts of the future. Risk is the possibility that a forecast will prove to be wrong. In the Penguin Dictionary of Economics, risk is defined as 'A state in which the number of possible future events exceeds the number of events that will actually occur, and some measure of probability can be attached to them' (Bannock et al 2003, p. 338).

A distinction is sometimes drawn between 'risk' and 'uncertainty'. Risk occurs where the probability distribution is known, and uncertainty where it is not. For the purposes of project evaluation, this distinction is irrelevant because any analysis of uncertainty requires a probability distribution to be specified. In the present context, the two terms may be used interchangeably.

An alternative approach in the literature is to define 'uncertainty' as imperfect knowledge about the future, and 'risk' as uncertain consequences. For example, the weather tomorrow is uncertain, but without consequences, there is no risk present. Risk will only exist when the weather has the potential to cause some economic loss or gain, physical damage or injury, or delay. (Austroads 2002, p. 3.)

The main sources of risk for public-sector projects are:

- construction costs that differ from expected because of changes in input costs or unforeseen events such as labour disputes or wet weather, or unforeseen technical factors. 'Scope creep', that is, increases in the scope of works to be undertaken as part of the project after evaluation has been completed, is a major source of construction cost overruns;
- operating costs that differ from expected because of changes in input costs or unforeseen technical factors;
- demand forecasts (and hence project benefits and operating costs) that differ from expected, a risk that rises the further into the future the projections are made;
- environmental impacts that differ from expected or were unforeseen;

- network effects, where an asset is part of the network (for example, an individual road) and decisions made elsewhere in the network impact on the project in question.

Risk premiums in cost–benefit analysis

Given a choice between a risky amount to be received in the future and an amount to be received with certainty, equal in size to the expected value of the risky amount, most people will prefer the certain amount. This is because they are 'risk averse', that is, risk is considered undesirable. For people to be indifferent between the two, the expected value of the risky amount must exceed the value of the certain amount. This difference is called the 'risk premium'. When comparing investments with differing degrees of risk, adjustments, based on risk premiums, have to be made to estimates of future returns. More risky returns have larger risk premiums and are therefore more heavily penalised. In the private sector, the necessary adjustments are made through applying a higher discount rate. Although a risk premium can be expressed as a simple dollar amount, it is more commonly viewed as the number of percentage points added to the risk-free discount rate to scale down future benefit or revenue streams to account for risk.

Downside risk

There is a good deal of evidence that **exante** evaluations of investment projects tend be over-optimistic when compared with **expost** performance.[1] People tend not to consider adequately what can go wrong, causing assessments to be biased in favour of the project. Also, where probability distributions are skewed, people choose the modal value of the variable (the highest point of the probability distribution) rather than the mean. Consequently, the forecasts employed in evaluations are more favourable than expected values (the mean of the probability distribution). For construction and operating costs, the tendency is to under-estimate, and for demand, to over-estimate. The difference between a projection biased on the optimistic side and the expected value can be termed 'downside risk'.

To correct the bias, there may be a preference for discount rates that are above the risk-free rate, incorporating a risk premium. However, except in rare circumstances, it has a distorting effect on project rankings.

1 See Flyvberg et al. (2003) for an extensive survey of the evidence.

Discounting reduces project benefits and costs, the further they occur in the future because the discount factor $1/(1+r)^t$ becomes smaller with time. The impact on project benefits of the addition of a risk premium to the discount rate therefore varies according to when the benefits occur in time.

In the absence of risk, the present value of a stream of benefits over the life of a project is given by:

$$\frac{B_1}{(1+r_o)^1} + \frac{B_2}{(1+r_o)^2} + \frac{B_3}{(1+r_o)^3} + \cdots + \frac{B_n}{(1+r_o)^n},$$

where r_o is the risk-free discount rate. The addition of a risk premium, r_p, to the discount rate can be represented as:

$$\frac{B_1}{(1+r_o+r_p)^1} + \frac{B_2}{(1+r_o+r_p)^2} + \frac{B_3}{(1+r_o+r_p)^3} + \cdots + \frac{B_n}{(1+r_o+r_p)^n}.$$

Thus benefits in each year are multiplied by a factor, F_t, of:

$$F_t = \left(\frac{1+r_o}{1+r_o+r_p}\right)^t \quad \text{where } t \text{ is the year.}$$

The factor will always be less than one and increases with time. For example, if the risk-free discount rate were 5 per cent and a risk premium of 2 per cent were added, then (using the formula just provided for F_t) first-year benefits would be multiplied by a factor of 0.98, second-year benefits by 0.96, and tenth-year benefits by 0.83.

The risk premium would properly adjust for downside risk only where the extent of over-estimation of net benefits rises each year in line with F_t. This is possible, but would occur only by coincidence. It should not occur where demand has been estimated using econometric modelling (regression of demand against variables such as income, prices and time). The statistical procedures typically used to make projections of demand growth are designed to minimise bias when applied correctly.

A risk premium will make no difference whatsoever to the present value of construction costs incurred in the year of analysis, that is, year zero for discounting purposes. Where planning, design and construction occurs over a number of years, only small percentage increases will be made to these costs as they are discounted forward to year zero, assuming year zero has been designated as the final year of construction. If year zero were designated as occurring at the **commencement** of construction, the risk premium would **reduce** construction costs incurred after the end of year zero.

Failure of net benefits to reach expectations may show up as soon as the project commences, not build up gradually over time in line with the impact of a risk premium. Examples would include operating costs being higher than expected and, in the case of a project that faces competition, a failure to attain the forecast market share.

A worrying possibility is that, if use of high discount rates to adjust for downside risk becomes the standard practice, analysts will be encouraged to make optimistic projections. Different studies may be carried out with different degrees of downside risk, and it is difficult to tell them apart. Applying the same risk premium across the board rewards projects with greater downside risk and penalises those with estimates closer to expected values. The danger is that the evaluation process will be reduced to one of project advocacy.

The risk premium approach breaks down completely where there are negative net returns in some years of a project's life. An example is a nuclear power station for which there is a large decommissioning cost at the end of its life. Raising the discount rate **reduces** the significance of future costs making the project appear **more** attractive rather than less.

In conclusion, raising the discount rate to incorporate a risk premium is a very poor way to correct for downside risk. The way to ensure that projections are free of downside risk is discussed below.

Pure risk

If downside risk has been eliminated from projections, there remains the variation about the expected value, called 'pure risk'. Can the results of CBAs be adjusted for pure risk by adding a risk premium onto the discount rate?

Before proceeding with the answer, a distinction has to be made between two types of pure risk: 'idiosycratic risk' and 'systematic risk'[2]. It is well known from portfolio theory that risk from share price movements that are uncorrelated with the share market as a whole can be eliminated by holding a diversified portfolio. Losses on some shares will be offset by gains on others, so the overall performance of the portfolio is smoothed. Idiosyncratic risk is that which can be removed by diversification. Systematic risk is that which cannot be diversified away. It arises from prices of individual shares being correlated with the price movements for the market as a whole.

2 'Idiosyncratic' risk is also called 'diversifiable' or 'non-systematic' risk. 'Systematic' risk is also called 'non-diversifiable risk'. The terms 'idiosyncratic' and 'systematic' risk have been used throughout this report to maintain consistency with Professor Quiggin's paper.

In the capital asset pricing model (CAPM), the level of systematic risk associated with a share determines the return demanded by the market above the risk-free return, that is, the risk premium. A share that has no correlation whatsoever with the market as a whole would be required to earn the risk-free rate of return, that is, zero risk premium. The higher the correlation, the greater the risk premium.

Idiosyncratic risk

Governments invest in large numbers of projects and their costs and benefits are spread over many individuals. Consequently, when individual projects over- or under-perform, the gains or losses are spread over many individuals, each of whom is only a little better or worse off than expected. Furthermore, the welfare of each individual will be affected by many projects. Just as for a share portfolio, the projects that do worse than expected will be more or less offset by projects that do better. So some of the risk associated with public-sector projects can be described as idiosyncratic. By definition, this risk can be entirely eliminated by diversification. Since individuals are not selecting a portfolio of public-sector projects in the same way they might select a share portfolio, some of the idiosyncratic risk will remain, but it is likely to be small enough so that it can be ignored for practical purposes.

Systematic risk

The systematic component of risk for public-sector projects arises from project benefits being correlated with benefits from other projects or with movements in the economy as a whole. To the extent that the benefits from a project are correlated with individual consumption, there are grounds for making a negative adjustment to a project's benefits.

As noted above, in the CAPM, it is systematic risk that gives rise to the risk premium. So the question arises: can the systematic risk associated with public-sector projects be incorporated into evaluations via a risk premium?

It was demonstrated above that adding a risk premium to the discount rate results in a set of downward adjustments to project benefits that increases over time. The set of adjustments will be the correct one to adjust for systematic risk only when the assumptions of the multi-period CAPM hold. These assumptions are likely to be met approximately for an investor considering share purchases but are more difficult for public-sector projects to meet. First, when comparing public-sector projects, a single risk premium applied across the board will be appropriate only if all the projects have similar risk characteristics. Otherwise, a different risk premium must be estimated for each project. Second, there has to be a single investment period followed by positive net benefit flows. The problems with applying risk premiums where construction occurs over several years and where there are years with negative net returns during the life of a project have already been discussed. Thirdly, the variance of the benefit stream has to increase linearly over time, which will occur only in special cases.

An approach to dealing with systematic risk for public-sector projects that does not involve adjustment of the discount rate and therefore does not rely on the restrictive assumptions of the multi-period CAPM is to estimate a certainty equivalent. The certainty equivalent concept is based on the 'expected utility' model. It explains why 'a bird in the hand is better than two in the bush'. Say you were offered a choice between receiving $5,000 with complete certainty or a 50:50 chance of receiving $10,000 or nothing. Even though the expected value of the risky option is $5,000 = 0.5 x $10000 + 0.5 x $0, most people would prefer the certain option. The reason, according to the expected utility model, is the 'diminishing marginal utility of money' — the more money you have, the less additional utility you gain from each additional dollar, and vice versa. So the total utility you gain from an extra $10,000 is something less than twice the total utility from an extra $5,000. The expected **utility**, as distinct from the expected dollar amount, from a 50:50 chance of receiving $10,000 is less than the utility to be had from $5,000 with certainty.

Say the amount you could receive with certainty was progressively reduced below the $5,000 starting amount until the point was reached where you were just willing to accept the 50:50 chance of receiving $10,000 or zero. The amount might be, say, $3,000. The expected utility from the 50:50 chance of receiving $10,000 or zero would be the same as the utility from receiving $3,000 with certainty. The $3,000 amount would be called the 'certainty equivalent'.

The 'risk premium' is the $2,000 difference between the expected value of $5,000 and the certainty equivalent of $3,000. The risk premium can be defined as the amount a person is willing to pay to avoid a risky situation. A diagrammatic exposition of the certainty equivalent and risk premium is provided in annex 1.

The risk premium is affected by the size of the uncertain amount relative to the individual's consumption. Most people would prefer a certain $5,000 to a 50:50 chance of $10,000 or nothing. The exceptions would be people who enjoyed gambling and/or were very wealthy. Expected utility theory cannot explain the joys of gambling. However, it is consistent with a wealthy person regarding plus or minus $5,000 as being of little consequence because they have a very low marginal utility of money, which would scarcely change over a range of $10,000. A person of modest means might feel similarly had the amount been $1 instead of $10,000.

If the individual's consumption is subject to fluctuations, then there is an additional factor to consider; the correlation between consumption and the risky amount to be received. Due to diminishing marginal utility of money, a sum of additional money will be worth more to an individual when consumption is low compared to when consumption is high. So if the amount of additional money were lower when consumption is low and higher when consumption is high, the value to the individual of the risky amount would be reduced; the certainty equivalent would be smaller and the risk premium higher. Conversely, if there were a negative correlation between consumption and the size of the risky amount to be received, the certainty equivalent could exceed the expected value of the risky amount resulting in a negative risk premium.

For example, say your annual income fluctuated with a 50:50 chance of being $30,000 or $60,000 in any year. In addition, each year there is a 50:50 chance that you will receive a further $10,000 or zero. Because of diminishing marginal utility of money, the additional $10,000 will be worth more in low-income years, and conversely. In this situation, it would be most desirable for the additional $10,000 to be perfectly negatively correlated with income, that is, you would receive the $10,000 in years when income was $30,000. The least desirable situation would be perfect positive correlation, with the $10,000 received in years when income was $60,000. Zero correlation, whereby receipt of the $10,000 was completely independent of income, would have an intermediate level of desirability. Derivation of the certainty equivalent and risk premium where there is perfect positive correlation is illustrated diagrammatically in annex 1.

The notion that the correlation between consumption and the risky amount to be received is an important determinant of the risk premium is a central conclusion of the 'consumption capital asset pricing model' (CCAPM). In the case of the **standard** CAPM, the expected rate of return on an individual share is related to the covariance between the return on the individual share and the return offered by the market as a whole. The return offered by the market as whole, over and above the risk-free rate, is exogenenous to the CAPM, and serves, in the model, as the price of risk. It is the role of the **consumption** CAPM to explain the return from the market as a whole. The CCAPM relates the market return to the covariance between the market return and the marginal utility of consumption, which in turn is a function of the level of consumption.

One of the assumptions underlying the mathematical representation of the CCAPM is that the risky amount to be received is small in relation to the individual's total consumption. The implication is that receipt of the risky amount of money leaves an individual's marginal utility of money practically unchanged. The risk premium arises because the marginal utility of money, and hence the value of the risky amount, changes with the level of consumption, combined with the existence of a correlation between consumption and the risky amount.

Annex 2 provides a mathematical derivation of the certainty equivalent and risk premium under assumptions of variable consumption and a relatively small amount at risk. In annex 3 the formula derived in annex 2 is used to make a general statement about the likely size of the risk premium for CBAs.

As shown in annex 3, under these assumptions, the risk premium, as a proportion of the risky amount to be received, depends on the product of four factors:

- a measure of the risk aversity of the individual (that is, the sensitivity of the marginal utility of money to changes in consumption);
- the variability (in proportional terms) of consumption over different states of nature (which affects the extent to which the marginal utility of money is likely to vary);
- the variability (in proportional terms) of the risky amount to be received; and
- the correlation coefficient between consumption and the risky amount to be received.

When the model is applied to public-sector investment projects, the conclusion is that the risk premium for systematic risk is likely to be very small in most cases. First, consider the variability of aggregate consumption over time. While growth rates for gross domestic product (GDP) fluctuate over time and occasionally become negative, the variation of GDP around the long-term trend amounts to no more than several percentage points. The same can be said for aggregate consumption, which is closely related to GDP. The variability of benefits from the typical project around its expected value that is correlated with aggregate consumption is also likely to be small. Multiplying these small amounts together produces a still smaller amount for the risk premium.

Making some reasonable assumptions about magnitudes, Professor Quiggin concludes that a typical risk premium would be of the order of 0.1 per cent. Thus a project with a BCR of over 1.001, evaluated without regard to risk, would still have a BCR of over 1.0 (a positive net present value) after adjusting for risk. Given the imprecision of CBA in general and of critical parameters such as the discount rate in particular, there is little to be gained from taking account of systematic risk. For further details on how this result was obtained, see annex 3.

The practical implication is that, if we attempted to adjust the results of CBAs to take account of systematic risk, then the adjustment would be so small as to be trivial, especially when compared with the margin for error. It might be noted that the adjustment would be made by reducing project benefits directly, not by raising the discount rate.

The conclusion that the risk premium for systematic risk is very small raises an interesting question. Long-term data from stock markets generally show that the rate of return from buying and holding the market portfolio of shares (all idiosyncratic risk diversified away) is considerably greater than the rate of return on government bonds (the risk-free rate). Yet the CCAPM predicts that the risk premium should be no more than half a per cent. It seems that individual investors are not behaving in the rational optimising manner that is the basic assumption underlying virtually all economic models. Economists refer to this apparent discrepancy between theory and real world behaviour as the 'equity premium puzzle'. A variety of explanations has been offered but there is as yet no consensus about a single one being right.

The explanations can be categorised according to whether or not they are consistent with the efficient markets hypothesis. If capital markets are efficient, then, at any given time, share prices fully reflect all available information. None of the explanations of the equity premium puzzle that are consistent with approximate market efficiency can be considered satisfactory.

It has been observed that the relative standard deviation of **individual** consumption is around 20 per cent, much greater than the 3 per cent variation for **aggregate** consumption. This implies that the extent to which risk in individual consumption is diversified away is considerably less than would be the case if the efficient market hypothesis were valid. Other explanations for the equity premium include:

- the possibility that investors over-estimate the riskiness of equity;
- the argument that investors prefer bonds to equity because bonds can be more readily converted back into cash;
- the absence of markets in which individuals can insure themselves against systematic risk in income from labour and non-corporate profits; and
- credit constraints or transactions costs associated with borrowing suppressing the demand for equity.

It may be that several of the explanations are correct and that the equity premium is the combined result.

Large effects on individuals

There is one case of pure risk that cannot be ignored. This is where a single investment project has a large effect on the welfare of a small number of individuals. Examples would include so-called essential services such as electricity, water supply and transport where the project represents the sole source of supply for a community.

The risk may be idiosyncratic or systematic or a mixture of both. To the extent that project benefits or costs are not correlated with the consumption of the individuals concerned, the risk is idiosyncratic. Since the effect on the welfare of the individuals concerned is large in relation to their total consumption levels, the idiosyncratic risk cannot be diversified away.

To the extent that project benefits or costs are correlated with consumption, the risk is systematic. Such a correlation could arise if the failures in the supply of an essential service were so pervasive that economic damage was done to the community in question reducing its overall consumption. The foregoing discussion based on the CCAPM does not apply because a major simplifying assumption, that the risky flow of income is small in relation to total consumption, does not hold. The argument that the risk premium is a small percentage of project benefits can therefore no longer be made.

Other arguments for high discount rates

BTE (1999) recommended discounting at the long-term government bond rate. Adding a risk premium is not the only argument for using a discount rate higher than the bond rate. This section addresses two of these other arguments:

- that the discount rate should be set at the social opportunity cost of capital; and
- that there is a need to adjust for the economic efficiency (also called 'deadweight') losses caused by higher taxes to fund projects.

Social opportunity cost of capital

When estimating costs in a CBA, the aim is always to measure the opportunity cost of resources used. Although the immediate source for funds for public-sector investment may be taxation, or charges levied for services, or borrowing, ultimately there are three sources of funds, each with its own opportunity cost:

- forgone private consumption;
- forgone private-sector investment; and
- borrowing from overseas sources.

The cost of funds for forgone private consumption is the social time preference rate (STP), the interest rate at which people are willing to trade-off present consumption for future consumption. For forgone private-sector investment, the cost of funds is the pre-tax rate of return that would be earned on the marginal private-sector project — the social opportunity cost (SOC). For overseas borrowing, the opportunity cost is the interest rate determined in international capital markets. In the absence of taxes, the costs of funds from all three sources would be identical and would equal the market rate of interest. Taxes on interest earned on peoples' savings drive down the social time preference rate. Taxes on corporate profits drive up the pre-tax rate of return required for private-sector investment. Hence the SOC of capital is well above the STP rate.

The SOC would be the appropriate discount rate to use only if all the funds required for investment in the project came from crowded-out private-sector investment, and if all benefits were invested in the private sector. In contrast, if all the funds came from forgone consumption and all benefits were consumed, the STP would be the appropriate discount rate. The same argument applies for overseas borrowing, with the real marginal cost of foreign borrowing used as the discount rate.[3]

Several different approaches have been proposed to account for the different opportunity costs of funds from different sources. The problem with applying them in practice is that they require knowledge of the STP rate and estimation of the shares of funds sourced from deferred private consumption and forgone investment. They may further require estimation of the destination of project benefits as between consumption and investment. Their data requirements make these approaches impractical, especially since share estimates are required for each individual project.

In the debate about discount rates, there is a general assumption of a closed economy. The third source of funds, overseas borrowing, tends to be ignored. For a small country with a good credit rating, the supply of international funds could be regarded as being perfectly elastic with respect to the interest rate over a large range. Of course, the level of borrowing cannot be increased indefinitely, without the interest rate rising. If all the benefits and costs of a project were to occur purely as changes in net foreign debt, the real interest rate on foreign debt would be the appropriate discount rate and it has the advantage of being readily measured.

While the cost of foreign borrowing may represent the true opportunity cost of funds **invested** in a public-sector project, there is still the destination of **benefits** to consider. When benefits take the form of increased consumption or investment rather than changes in borrowing, the problem again arises of the prohibitive data requirements of estimating shares of benefits going to the different destinations and the STP rate. Thus even with the assumptions of an open economy and a perfect elastic supply of funds, there is no practical solution to the discounting problem. (BTE 1999, pp. 70–71)

In the absence of a better solution, BTE (1999, p. 78) concluded that the most appropriate discount rate to use for CBA is the government bond rate. Because the government will not default on loans, the bond rate provides a ready measure of the cost of capital free of any risk premium.

3 The real cost of foreign borrowing would include front-end and commitment fees and risk premiums paid to foreign lenders. The reason for including risk premiums paid to foreign lenders is that CBA is normally undertaken from a national viewpoint. Payment of the risk premium to foreign lenders represents a loss of resources to the nation. For further discussion on how the real cost of foreign borrowing can be determined, see ADB 2005, chapter XI.

Deadweight loss of taxation

Another argument for using a discount rate above the government bond rate is to take account of the deadweight losses imposed on society by increased taxation to fund government investment projects. Estimates of the marginal deadweight loss from a tax increase vary widely even for a single given tax. BTE (1999, pp. 82–84) cites estimates ranging from $0.11 to $0.65 for an additional dollar raised from increasing taxes on labour income, and as high as $1.31 for taxes on spirits.

Just as adding a risk premium to the discount rate is a very poor way to account for risk, an increased discount rate is a poor way to adjust for the deadweight losses arising from taxation. As shown earlier, a higher discount rate adjusts the stream of project costs and benefits according to a particular pattern over time. If the tax increase to fund a project occurs during the construction phase of the project, the deadweight losses will accrue at that time. Yet a higher discount rate has little or no impact on investment costs and penalises benefits.

To be consistent, project benefits should be adjusted **upward** to the extent that they lead to increases in government revenue making tax reductions possible. Project benefits could lead to higher government revenues directly via charges for the project's services and indirectly if they stimulate economic activity leading to increased tax collections.

Borrowing to fund the project could spread the tax increases over time as the principal and interest on the loan were paid. This would reduce the deadweight losses because the marginal deadweight loss from a tax increase rises more than proportionately with the revenue raised. However, the pattern of deadweight losses over time is most unlikely to mirror the negative adjustments to project benefits stemming from a higher discount rate.

Furthermore, as BTE 1999 (pp. 80-82) points out, increased taxes are not the only source of funds for government investments. Levying charges on project beneficiaries and reducing other forms of government spending are alternatives that do not necessitate tax increases.

In practice, government spending decisions are made within budget constraints, so it is not a case of implementing all projects with BCRs above unity. With a constrained budget, a decision to implement any one project means non-implementation of other projects, not increased taxation.

In a budget-constrained situation, economically efficient selection of projects requires that projects be ranked in descending order of BCR. Projects with BCRs below the BCR of the last project chosen before the available funds were exhausted (the cut-off BCR), would not be implemented. Say the cut-off BCR was 3.0. Only projects with a BCR above 3.0 would be implemented. The outcome would be exactly the same as for a decision rule under which investment costs (the denominator of the BCR) for all projects were multiplied by a factor of 3.0 and all projects with a BCR greater than unity were implemented. It would be as though the government had determined that the opportunity cost of a dollar of government investment spending was three dollars. Hence the imposition of a budget constraint on government investment spending has the same effect as making an across-the-board upward adjustment to the investment costs of all projects under consideration during the budget period.

Assuming away any impacts on government revenues arising from project benefits and costs following the construction period, the optimal amount of tax to raise to pay for a government investment would be that at which the marginal deadweight loss from taxation (expressed as a ratio) equates to the cut-off BCR minus one. For example, say raising an additional dollar of tax imposed a deadweight loss of one dollar (a ratio of 1.0) and the cut-off BCR was 3.0. Then by raising an additional dollar in tax and investing it, society would gain $2 in net benefit at the expense of $1 in deadweight loss from taxation — a net gain of a dollar. As taxes were raised to fund additional government investment, the marginal deadweight loss from taxation would rise and the cut-off BCR would fall, until the optimum was reached. If the point was reached where the marginal deadweight loss ratio was 1.5 ($1.50 for each additional dollar raised from taxation) and the cut-off BCR as 2.5 ($1.50 of net benefit for each additional dollar invested), no further gains could be made by changing the levels of taxation and investment.

Or course, the factors that influence government decisions about levels of taxes and investment spending are much broader and more complex than this suggests. However, it could be argued that, at least in principal, where total investment spending is constrained by a budget, the deadweight loss of taxation is implicitly addressed through the government's budgetary processes.

Given that the primary taxation impacts of projects arise from capital costs incurred during the construction phase rather than subsequent costs and benefits, budget constraints on government investment are arguably a better way to address deadweight losses from taxation than higher discount rates.

What to do in practice

It has been argued that adding a risk premium to the discount rate can only be justified under assumptions that are unlikely to hold for most applications of CBA. Adjustments to account for risk need to be made directly to project benefits and costs.

The practical implications for dealing with the different types of risk can be summed up as follows:

- downside risk: attempt to minimise it by ensuring that the costs and benefits in CBAs are expected values;
- pure risk (both idiosyncratic and systematic): the required adjustments are usually so small that, for practical purposes, it can be ignored. The exception is the case where the project has potentially large effects on the welfare of affected individuals.

This section explains in detail how to minimise downside risk and how to estimate a certainty equivalent where required.

Treating downside risk

Downside risk arises, in the main, from a failure to consider what can go wrong. So if downside risk is to be minimised, a complete assessment needs to be made of the possibilities. The 'state-contingent' approach provides a framework for doing this. It represents a thought process that disciplines the analyst to ask a complete set of 'what if' questions.

Some definitions

Outcome

An **outcome** is the result of a situation involving uncertainty. For example, a project failing for technical reasons is an outcome. Absence of failure for technical reasons is another outcome. For forecast demand, an average annual daily traffic level (AADT) of 11,841.4 vehicles per day is an outcome[4]. The former is a discrete variable and the latter is continuous.

Event

An **event** is any collection of outcomes. For example, partial failure and complete failure on technical grounds could be grouped together and defined as an event, 'partial failure or worse'. For a continuous variable, an AADT in the range >8000 to ≤12000 could be defined as an event.

Event space

An **event space** (also known in statistics as a sample space or possibility space) is the set of all the possible events. For technical success of the project the event space might be {fail, partial fail, not fail} or {partial failure or worse, not fail} Outcomes or events may be distinguished by time, for example {fail in year 1, fail in year 2, fail in year 3, ..., fail in year 20, not fail}.

Since the list of outcomes in an event space is exhaustive, the probabilities of all events in an event space should sum to one. For example, if the event space for the technical success or failure of a piece of infrastructure were {fail, not fail}, the probabilities could be {0.01, 0.99}.

For a continuous variable, such as AADT, the event space could range from zero to infinity. The possibilities for partitioning the event space for a continuous variable into events are limitless. For AADTs, examples would be {≤10000, >10000} and {≤8000, >8000 to ≤12000, >12000}. A continuous variable will have a probability distribution associated with it. The probability associated with each event in a continuous event space would be measured as the area under the probability distribution between the event boundaries.

State of nature

A **state of nature** is a collection of events selected from different event spaces. For example, the project not failing **and** AADT being greater than 12,000, would be a state of nature.

State space

A **state space** is the set of all possible states of nature. If there are n event spaces identified by the subscripts $i=1, 2, 3, \ldots, n$, and each event space contains m_i outcomes, the maximum number of possible states of nature will be $m_1 \times m_2 \times m_3 \times \ldots \times m_n$. The reason why this is a maximum and not a total is that some events make the events in other event spaces redundant. For example, if the project is a road tunnel and it fails altogether, the different possible AADT

4 Average annual daily traffic (AADT) is the number of vehicles passing a given point on a road during a year divided by the number of days in the year.

outcomes are irrelevant. If one or more event spaces are treated as continuous variables without partitioning into discrete events, the number of states of nature becomes infinite.

The probability associated with a particular state of nature is the product of the probabilities of its constituent events.

The probabilities of all the possible states of nature must sum to one.

Event trees

States of nature may be identified using an event tree.

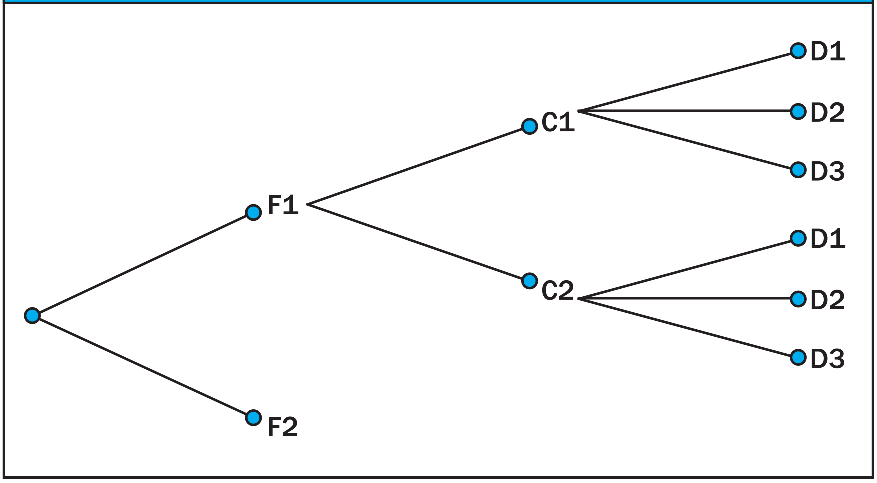
Say for a transport infrastructure project, there are three event spaces (represented as F, C and D) comprised of events as follows.[5] The individual events comprising the event spaces are represented as F1, F2, C1, C2, D1, D2 and D3.

- The project may proceed normally (F1) or fail due to unexpected technical difficulties and have to be abandoned (F2);
- annual operating costs may be $10m (C1) or $20m (C2); and
- annual growth in demand may be 2 per cent (D1), 4 per cent (D2) or zero (D3).

There are 2x2x3 = 12 combinations, but only seven states of nature, because failure of the project (F2) renders the other event spaces irrelevant. They might be mapped out in an event tree as in figure 2. Events that render other events redundant should be placed in front of or above the events they make redundant.

5 This example consists of the same events and probabilities as example 2 in Professor Quiggin's paper. The cost and benefit values, however, have been changed.

FIGURE 2 AN EVENT TREE



The next step is to assign a probability to each occurrence of each event. For example, the probability of F2, technical failure, might be 0.2 and F1, success 0.8. The event tree can be set out in tabular form as shown in table 1, with probabilities attached. In practice, there will often be a fair amount of subjectivity in estimating probabilities. Sometimes, historical data or engineering models can assist.

Table 1 An event tree in tabular form

Technical failure		First year costs		Demand growth		Probability of state of nature
Event	*Probability*	*Event*	*Probability*	*Event*	*Probability*	
F1	0.8	C1	0.6	D1	0.6	0.288
				D2	0.2	0.096
				D3	0.2	0.096
		C2	0.4	D1	0.6	0.192
				D2	0.2	0.064
				D3	0.2	0.064
F2	0.2					0.200
Total						1.000

The level of any benefit or cost is contingent upon a state of nature. Having specified the states of nature and probabilities, the values of benefits or costs have then to be estimated for each year of the project's life in each state of nature.

A CBA is always a comparison between two states of the world, a base case and a project case. Normally the base case is the situation where the project does not proceed. A benefit or cost is the difference between the forecast level of a variable in the base case and its forecast level in the project case. It should be noted that the **base case** could be different in different states of nature as well as the project case. For example, flooding could lead to rapid deterioration of an **existing** road in the base case. The benefit from replacing it with a new road that is less vulnerable to flood damage would therefore be greater in states of nature in which flooding occurs because the base case will be worse.

Table 2 completes the numerical example. Other project assumptions made to develop the list of annual costs and benefits under each state of nature are that:

- the construction cost is $100m spread evenly over two years;
- project life is 8 years after completion of construction;
- benefits are $40m in the first year and grow at the same rate as demand; and
- the discount rate is 5 per cent.

Table 2 Calculation of expected NPV and BCR

($ millions)

	C1D1	*C1D2*	*C1D3*	*C2D1*	*C2D2*	*C2D3*	*F2*	*Expected values*
Year	Annual net benefits							Mean
0	-50	-50	-50	-50	-50	-50	-50	-50
1	-50	-50	-50	-50	-50	-50	-50	-50
2	30	30	30	20	20	20	0	21
3	31	32	30	21	22	20	0	21
4	32	33	30	22	23	20	0	22
5	32	35	30	22	25	20	0	23
6	33	37	30	23	27	20	0	23
7	34	39	30	24	29	20	0	24
8	35	41	30	25	31	20	0	25
9	36	43	30	26	33	20	0	26
Probability	0.288	0.096	0.096	0.192	0.064	0.064	0.200	
NPV	104	122	87	42	60	25	-98	44
NPV x prob	29.9	11.7	8.4	8.1	3.8	1.6	-19.5	44
PV benefits	201	219	185	140	158	123	0	
PV costs	98	98	98	98	98	98	98	
BCR	2.1	2.2	1.9	1.4	1.6	1.3	0.0	
BCR x prob	0.59	0.22	0.18	0.27	0.10	0.08	0.00	1.4
IRR	23.3%	25.3%	21.3%	13.2%	15.8%	10.4%	-200.0%	

Note: Totals may not add due to rounding.

Expected value of net present value

The final step is to calculate the expected value of each benefit and cost for each year using the probabilities, and to discount them at the risk-free rate. All costs and benefits for a state of nature need to be multiplied by the probability for that state and the results summed. Alternatively, one could discount first to obtain the net present value under each state of nature, and then derive the expected net present value by multiplying by probabilities and summing. Table 2 shows that the same result, $44m, is obtained both ways.

Expected values of other CBA results

The NPV is used to determine whether or not a project is economically warranted and so whether it should proceed. It is also used for comparing mutually exclusive projects, for example, different route options for a road or different sizes or implementation times for the same project. Results of CBAs are often presented as BCRs or internal rates of return (IRRs). Multiplying by probabilities before or after combining costs and benefits is **not** a matter of choice for calculating expected values of BCRs or IRRs. Before going on to demonstrate why this is the case, the BCR and the IRR concepts and their uses are explained.

The BCR is used for ranking projects where there is budget constraint. It is also useful for presenting results of CBAs in a way that can be easily understood. The formula is:

$$\text{BCR} = \frac{\text{Present value of [benefits minus operating costs]}}{\text{Present value of investment costs}}$$

The project's operating costs are treated as a negative benefit and appear in the numerator and the investment costs appear in the denominator because it is assumed that only investment costs are being rationed. For a given budget out of which investment costs are paid, undertaking the projects with the highest BCRs will deliver the largest total of net benefits achievable within the budget constraint.

If operating costs are significant and have to be paid out of future budgets so that they will have to compete for funds with investment in future projects, the situation is more complicated. An optimisation problem could to be set up with assumptions made about the sizes of future budgets and the benefits and costs of future projects. A simple approach is to assume values for cut-off BCRs in future years. Operating costs for all projects being ranked in the current period could then be multiplied by the cut-off BCR in the year in which they will be incurred. For example, say a project implemented today creates a maintenance need of $1 million in 10 years time that will have to be funded out the same budget as for capital projects. Assuming that the cut-off BCR in 10 years time is 3.0, spending $1 million on maintenance in 10 years time will preclude the opportunity to invest $1 million in capital projects with a BCR of 3.0. Hence the opportunity cost of the maintenance commitment generated by the current project is $3 million in forgone benefits in 10 years time. For a detailed discussion of the issue see ATC (2004, volume 3, pp. 65–67).

The IRR is the discount rate at which the NPV equals zero. It can be interpreted as the minimum value of the discount rate at which the project is economically warranted. Like the BCR, the IRR presents the results of a CBA in a way that is easy for people to understand. It has the added advantage that it obviates the need to be specific about the discount rate. The project is economically justifiable if the IRR exceeds the discount rate. Projects should never be ranked or compared using internal rates of return. This would be equivalent to evaluating projects at different discount rates. Where the IRR is significantly different from the correct discount rate, the results could be quite misleading.

In table 2, for the BCR, the expected value is 1.4, obtained by taking the expectation of the BCRs for each state of nature. The expected IRR is problematic because the IRR for the F2 state of nature of negative 200 per cent — loss of the entire capital with no return — distorts the result. The IRR rate of return measure cannot be relied upon when the pattern of costs and benefits over time departs from the conventional one of early costs during the investment phase, followed by a continuous stream of positive benefits. So the expected IRR is not available in this case.

The choice of estimating expected values before or after combining costs and benefits is not available for the BCR and IRR measures. Multiplying all costs and benefits for a state of nature by a probability has no effect whatsoever on the BCR or the IRR. These measures must be calculated for each state of nature **first**, and only then can the expected values be derived.

The numerical example in tables 3 and 4 demonstrates this. Assume that a project has:

- an investment cost of $100 million;
- a life of three years;
- annual benefits of $80 million in state of nature A and $50 million in state of nature B with each state of nature having a 50 percent probability; and
- annual operating costs of $10 million.

Investment and operating costs are the same in both states of nature. The discount rate is 5 per cent.

Table 3 shows the CBA performed under each of the two states of nature, with the expected values of the results calculated at the end of the process. With a 50 per cent probability for each state of nature, the expected values of the NPV, BCR and IRR fall midway between the values estimated for each state of nature.

In table 4, the costs and benefits have been multiplied by the probabilities **before** discounting. With costs and benefits multiplied by 0.5, the NPV under each state of nature is halved, but the BCR and IRR are unchanged. When the results are added together at the end of the process to estimate expected values, the NPV is correct, but the BCR and IRR are twice the correct results.

Table 3 Estimation of expected values of cost–benefit analysis: multiplication by probablilities at the end of the process

($ millions)

	State of nature A				
Investment cost	100	100			
Benefits minus operating costs	191		70	70	70
Net present value	91				
Benefit–cost ratio	1.9				
Internal rate of return	48.7%				
	State of nature B				
Investment cost	100	100			
Benefits minus operating costs	109		40	40	40
Net present value	9				
Benefit–cost ratio	1.1				
Internal rate of return	9.7%				
	Expected values of results				
Net present value	50				
Benefit–cost ratio	1.5				
Internal rate of return	29.2%				

Table 4 Estimation of expected values of cost–benefit analysis: multiplication by probabilities before discounting

($ millions)

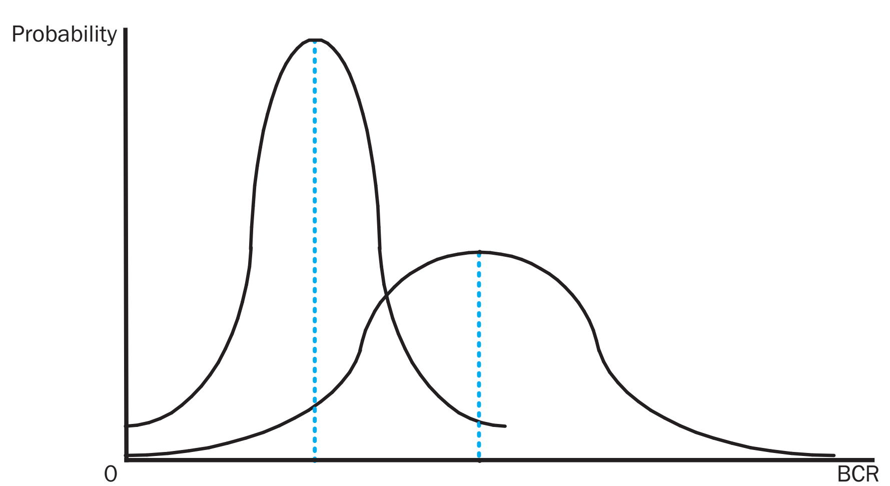
	State of nature A				
Investment cost	50	50			
Benefits minus operating costs	95		35	35	35
Net present value	45				
Benefit–cost ratio	1.9				
Internal rate of return	48.7%				
	State of nature B				
Investment cost	50	50			
Benefits minus operating costs	54		20	20	20
Net present value	5				
Benefit–cost ratio	1.1				
Internal rate of return	9.7%				
	Expected values of results				
Net present value	50				
Benefit–cost ratio	3.0				
Internal rate of return	58.4%				

Computer software

Where the number of states of nature and/or the number of uncertain variables are large, the number of combinations of input values can become extremely large. A computer program such as @RISK, which links with widely-used spreadsheet programs such as Excel and Lotus 123, can facilitate the process. Such programs use sampling procedures with the number of iterations specified by the user. For continuous variables, the user can specify probability distributions. In a typical iteration, the program would draw a random sample value for each uncertain variable, with the sampling dictated by the associated probability distributions. The set of sample values would be inserted into the CBA spreadsheet to find the NPV, BCR and other required results. Repeating the process a large number of times, the program constructs probability distributions of CBA results.

From the probability distributions, estimates of the variances of results are available as well as expected values. The variance of a CBA result is an indicator of pure risk, though not a complete indicator because it provides no information about the correlation with the general level of economic activity. Given that pure risk can be ignored in most cases, the variance would normally be of no relevance whatsoever to decisions about whether projects should proceed or for ranking projects. Except in the case where projects have potentially large impacts on the welfare of a small number of individuals, one project would not receive preference over another with a higher economic return on the grounds that the former has a smaller variance. In figure 3, the project with the higher expected BCR is still preferred even though it has a greater variance.

FIGURE 3 COMPARING PROBABILITY DISTRIBUTIONS OF BENEFIT-COST RATIOS



For more discussion on the use of @RISK in the road project evaluation context, including a worked example, see Austroads (2002).

Optimal timing

In the absence of risk, the rule for optimal timing of investment projects is that deferral may be warranted if the first-year rate-of-return (FYRR) criterion is not met. Provided project net benefits grow over the life of the project, the optimal implementation time is the first year in which the FYRR exceeds the discount rate, that is:

$$\frac{B_1}{K} > r$$

where B_1 is net benefits in the first year of the project's life and K is the capital cost. If the project is delayed by one year, society forgoes B_1 in benefits, but gains rK, the time value of deferring the capital cost by one year. If $B_1<rK$, then society gains more by deferring the project than it loses. As demand grows over time, first-year benefits rise until the point is reached when $B_1>rK$, that is, the benefit lost by delaying the project another year is greater than the capital cost saved. For a diagrammatical exposition and a mathematical derivation of the FYRR criterion, see annex 4.

If net benefits decline at any stage of the life the project, the criterion could indicate more than one possible optimum implementation time. The reason is explained in annex 4. In such cases, NPVs should be calculated for each possible optimum time in order to find the one that yields the maximum NPV.

The first-year rate-of-return criterion can be applied **in the presence of risk** using the expected values of the capital cost and first-year benefit.

Changes in project implementation time can give rise to changes in probabilities and magnitudes of costs and benefits. Where expected net benefits or capital costs vary with implementation time, it would be prudent check for other possible optimum times, using the first-year rate-of-return criterion, and to compare their NPVs. The case where project deferment alters the probabilities of states of nature is discussed further below.

It should be noted that the NPV calculations have all to be made from the same year of analysis. For example, if the present were chosen as the year of analysis, implementation in five years time would cause construction costs to be discounted over five years, first-year benefits to discounted over six years and so on. Implementation in year 10 would cause construction costs to be discounted over 10 years, first-year benefits over 11 years and so on.

The discounting period has to be extended far enough into the future so that changing the final year of discounting, as the project is delayed, has negligible effect on the net present value. The lower the discount rate is in relation to the annual rate growth rate for benefits, the longer the required time period.

Estimating a certainty equivalent

A certainty equivalent of a benefit or cost may be required where the welfare of a small number of individuals affected by the project varies greatly across the states of nature. The simplest practical approach to estimating a certainty equivalent involves assuming a utility function from which to estimate the level of expected utility and then to convert back from utility to dollars.

The state-contingent approach just described should be followed, specifying states of nature and estimating the probability and values of costs and benefits for each state. For the group of individuals whose welfare is subject to great variability, their consumption levels in the base case and project case, in each state of nature should be estimated. Then:

- assume a utility function;
- use the utility function to convert base-case and project-case dollar values of consumption to utility for each state of nature;
- calculate the expected utility across all states of nature for the base case and project case;
- convert the base-case and project-case expected utilities to dollar values using the utility function; and
- take the difference between base-case and project-case dollar values to obtain the value of the benefit or cost to the individuals affected.

This will have to be done for each year separately. The resultant values of the costs and benefits should be inserted into the analysis **before** any discounting is undertaken.

If there are large disparities in income levels within the group, a better estimate will be obtained by segmenting the group according to income level and making separate estimates of utility levels, benefits and costs for each sub-group.

A suitable form of utility function to assume is:

$$u(c) = \frac{kc^{1-\gamma}}{1-\gamma} \text{ for } \gamma \neq 1 \text{ or}$$

$$u(c) = k\,ln(c) \text{ for } \gamma = 1.$$

Most studies suggest that $0 \leq \gamma \leq 2$. The constant γ is called the 'coefficient of relative risk aversion'. A value of zero implies risk neutrality. Higher values above zero indicate greater levels of risk aversity. *k* is a scaling factor and may be set equal to $(1-\gamma)c_0^{\gamma}$ for $\gamma \neq 1$ or $c_0/\ln(c_0)$ for $\gamma = 1$ where c_0 is base-level consumption. At consumption level c_0, one unit of utility will be worth exactly one dollar.

In the absence of better information, it would be reasonable to assume γ is unity and perhaps do sensitivity tests for values of 0.5 and 1.5, or zero and 2.0.

A simple worked example is provided in annex 5.

Implications for risk management

Risk management can be defined as the process of assessing exposure to risk and determining how best to handle such exposure with the aims of minimising risk and optimising the risk–benefit balance. The conclusion that pure risk can reasonably be ignored in most situations has implications for risk management. Alternative risk management strategies can be compared using the state-contingent approach to find the one that yields the highest expected NPV. This greatly simplifies comparisons between project options having different levels of risks and costs. There is no need to estimate certainty equivalents with the requirement to make subjective judgements about the degree of risk aversity (curvature of the utility function).

In the numerical example of table 1, demonstrating the state-contingent approach, there was a 20 per cent probability of technical failure of the project. In practice, such a large chance of technical failure would be unacceptable for a major infrastructure project. Ways would be found to reduce the probability of failure even though they increased the cost of the project. For example, a bridge could be built to have greater strength, or, if the project was a tunnel, more extensive geological studies could be undertaken to better understand the risks and to devise actions to reduce the probability of major technical difficulties arising during construction.

Risk management is usually thought of in terms of undesirable outcomes, both in terms of the probabilities and the costs imposed. However, it is possible for a project to be over-designed, — that is, it is preferable to accept an increased risk of an undesirable outcome in exchange for a cost saving.

The twin aims of risk management are risk minimisation and optimising the risk–benefit balance. Risk minimisation means taking any actions that reduce the probability or the cost of undesirable outcomes, where these actions involve little or no cost. Risks also have to be minimised where required by law or by community expectations. Attaining the optimal risk–benefit balance requires determination of how much net benefit to sacrifice in order to reduce risk. In most cases, attaining the optimal risk–benefit balance is a matter of determining the option with the highest expected NPV.

Table 5 shows two project options with different risk–benefit trade-offs. Spending an additional $50m reduces the chance of technical failure from 20 per cent to 5 per cent, generates a net gain in expected NPV of $25m after reducing the NPVs for both states of nature by $50m. The gain comes about because of the change in probabilities.

Table 5 Comparison of risk options using expected NPV: changing probabilities

($ millions)

Without extra spending to reduce risk				
Pass	$200	0.8	$160	
Fail	–$300	0.2	–$60	
				$100
With an extra $50m spent to reduce risk				
Pass	$150	0.95	$142.5	
Fail	–$350	0.05	–$17.5	
				$125

Another option might leave the probabilities unchanged, but reduce the cost of failure. Such would be the case for measures taken to lessen damage in the event of a disaster. In table 6, spending an additional $50m reduces the NVP of failure to zero, so that the NPV in this state of nature becomes –$50m after allowing for the additional spending. The net benefit is an increase of $10m in the expected NPV. The net gain from spending to reduce the cost of failure is given by:

(probability of failure x present value of benefit) – present value of cost

This amount has to be positive for the investment to be worthwhile. The benefit, that is, the reduction in cost to society in the event of failure, is only realised in the event of failure and so is multiplied by the probability. The cost is incurred in all states of nature and so is multiplied by unity. In terms of the example in table 6, the net benefit is $10m = 0.2 x $300m – $50m.

Table 6 Comparison of risk options using expected NPV: reducing costs of failure

($ millions)

Without extra spending to reduce cost of failure				
Pass	$200	0.8	$160	
Fail	–$300	0.2	–$60	
				$100
With an extra $50m spent to reduce cost of failure				
Pass	$150	0.8	$120	
Fail	–$50	0.2	–$10	
				$110

Risk–benefit trade-offs would have to be considered in cases where the certainty equivalent is required, that is, where the particular risks could be associated with large changes in the welfare of a small number of individuals. In these rare situations, the objective would be to maximise the certainty equivalent of the NPV.

Project deferment

Project deferment is one of the many strategies available for managing risk.

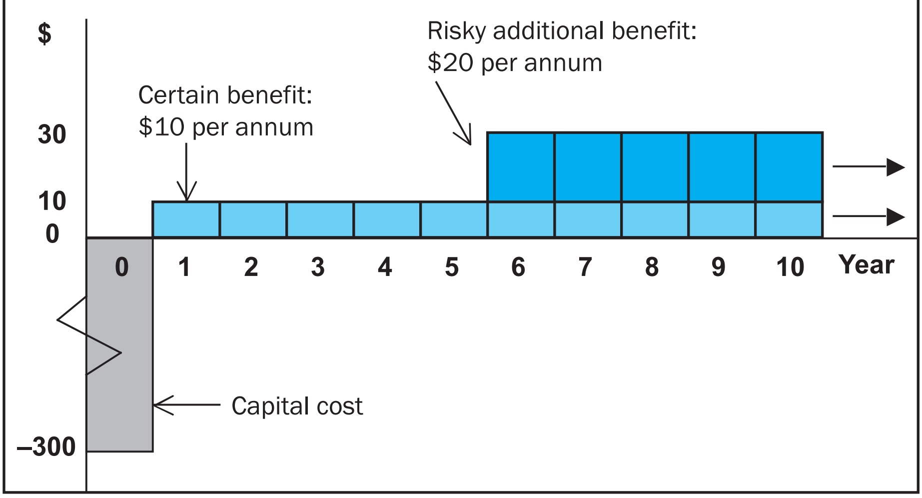
Probabilities of states of nature can change with implementation time in situations where the costs or benefits of a project are affected by an uncertain one-off event in the near future. The option may exist to defer the project until after the outcome of the one-off event is known and the uncertainty has ceased to exist. The possibility of a negative outcome for the project is thereby avoided. A wait-and-see policy might be worthwhile when project benefits or costs are highly dependent on a future decision by government to authorise or by the private sector to undertake a significant residential or industrial development. Other uncertain future events that could be worth waiting for include decisions to proceed with other projects in a network (network risk) and possible significant changes in the overall level of economic activity or in exchange rates.

The wait-and-see option will be preferable if it is associated with a higher expected NPV than the build-now option, calculated from the standpoint of a common year of analysis.

BTE (1999, pp. 75–6) uses the example of construction of a new highway where there are alternative routes. The benefits from the various routes depend on patterns of regional growth. If construction of the highway is deferred until such time as regional growth patterns have become clearer, the benefits of each route option can be evaluated with greater certainty. The option of not constructing the highway at all is also available.

The following simple numerical example provides an illustration. Assume that a freeway costs \$300m to build and will yield annual benefits of \$10m with certainty. Figure 4 shows the flows of costs and benefits. For simplicity, assume that annual benefits are constant over time and continue forever. Hence, at a 5 per cent discount rate, the present value of benefits is \$200m = \$10m/0.05. The NPV is therefore –\$100 = \$200m – \$300m. There is a 50 per cent probability that a new satellite city will grow up generating additional traffic, giving rise to a further \$20m per annum in benefits commencing in 5 years time. Under this state of nature, the present value of annual benefits (\$10m for the first 5 years and \$30m in perpetuity thereafter) is \$513m, implying an NPV of \$213m = \$513m – \$300m. With a 50 per cent probability of the satellite city growing, the expected NPV is \$57m = 0.5 x (–\$100m) + 0.5 x \$213m.

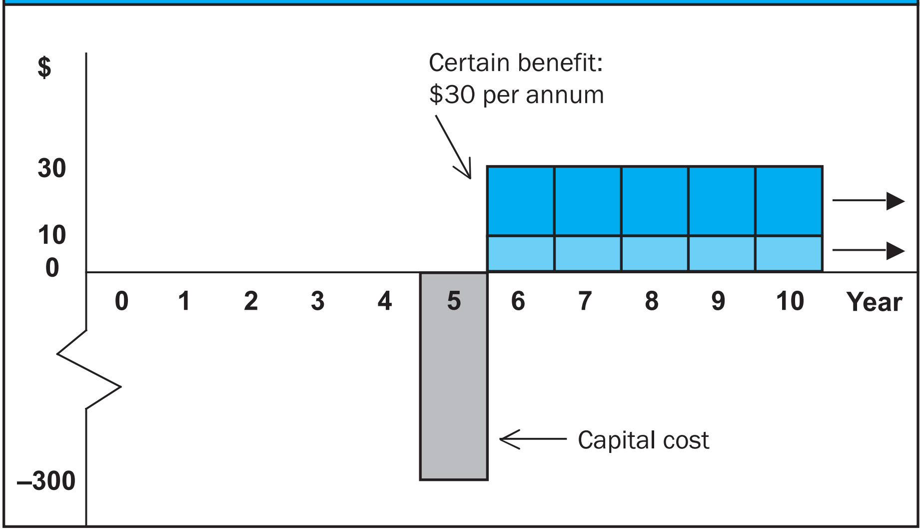
FIGURE 4 ILLUSTRATIVE PROJECT CAPITAL COST AND BENEFITS



However, there is another option to consider. If the project is delayed by 5 years, it will be evident by then whether or not the satellite city has grown up. The risk will have vanished. In order to compare the deferral option with the build-now option, the costs and benefits of the deferral case have to be discounted to today (year zero). Figure 5 illustrates the stream of costs and benefits. The benefit stream is zero for years one to five and \$30m in perpetuity thereafter, yielding a present value of \$470m. By deferring the project, society has forgone the \$10m of certain benefits for the first 5 years. Offsetting this, deferral of capital costs reduces the capital cost in today's dollars to \$235m = \$300m/1.05^5. The NPV is \$235m = \$470m – \$235m, which would be gained with certainty once it is known that the satellite city will grow. If the satellite city has not grown up in 5 years time, the decision will be not to proceed with the project and the NPV will be zero.

From the standpoint of today, each outcome still has a 50 per cent probability. Hence, the expected NPV under the deferral option is \$117.5m = 0.5 x \$235m + 0.5 x \$0m. This is \$61m greater than the expected NPV of investing now, and so is preferable. By deferring the project, the 50 per cent chance of a loss to society of \$100m has been eliminated.

FIGURE 5 ILLUSTRATIVE PROJECT CAPITAL COST AND BENEFITS: PROJECT DEFERRED FIVE YEARS



The deferral option may impose some additional costs not incurred under the build-now option. For example, if land has to be reserved to maintain the option, the land may have to be purchased now instead of in five years time. A cheaper alternative might be to pay the owner an amount in exchange for an option to purchase land in five years time. The amount of the payment would not enter into the CBA because it is a transfer, but any costs associated with the restricted use of the land because of the risk of resumption in five years time would be counted. Costs of keeping the deferral option open should be built into the calculation.

Another way to view the situation parallels the 'real options' approach used in finance. It shows how to derive a monetary value for the option of deferral. The \$61m expected gain from deferral could be considered as the value to society of having the option to defer available. If there were resource costs involved in preserving the option, \$61m would be the upper limit on the amount worth incurring.

For example, if preservation of the option necessitated purchase of the land now, holding it for 5 years, the option cost would be the present value of the time cost of the land for the 5-year period[6]. A less expensive alternative could be to offer the owner of the land a sum of money in exchange for an option to purchase the land in 5 years time at either an agreed fixed price or at the future market price. It should be noted that, for CBA purposes, the cost to society is not the sum paid to the owner, because it is a transfer payment. The relevant cost for the CBA is the resource cost of the restriction in use of the land created by the 50 per cent chance of it being resumed for a freeway in 5 years time.[7]

For further discussion and references concerning real options, see BTE (1999, pp. 75–76), Brealey and Myers (2003), and Luehrman (1998).

An important point to note from the discussions of optimal timing and project deferral is that a positive expected NPV does not necessarily imply that a project should proceed. In addition to testing alternative options for the project itself to find the one that maximises the NPV, different timing options can be tested taking account of any changes in probabilities of states of nature as the implementation time changes, and of the costs of keeping options open when projects are deferred.

A related risk management strategy is to divide the project into stages. It can reduce risks by deferring commitment of part of the total investment until uncertain developments have played out. Also, better market information can be obtained by 'testing the waters' before committing to the entire investment. Staged options can be evaluated in the same way as for other options with differing risks, by estimating expected NPVs.

---

6 The cost of purchasing the land now instead of in 5 years time is $C - \frac{C}{(1+r)^5}$ where C is the cost of the land. This is equal to the present value of the interest forgone:

$$\frac{rC}{(1+r)} + \frac{rC}{(1+r)^2} + \frac{rC}{(1+r)^3} + \frac{rC}{(1+r)^4} + \frac{rC}{(1+r)^5}$$

7 The cost of restricted use of the land would be built into the option price, but there are other factors affecting the option price. Distortions in prices or externalities could give rise to differences between private and resource costs. If the option involves purchase of the land at a fixed price, the option price will be affected by the level of the agreed land price in relation to the current market price, and expectations about future market prices. The relative bargaining abilities of the parties involved could also influence the outcome.

Implications for public versus private sector investment

Discounting public-sector projects at the risk-free rate, that is, the long-term bond rate, creates an appearance of giving preferential treatment to public-sector projects over private-sector projects. It might be thought that it would lead to higher levels of government investment crowding out some private investment. Before examining the issue in detail, the treatment of risk in financial analysis is briefly discussed.

Discounting and risk in financial analysis

Just as for a CBA, for a financial analysis, the discount rate should be chosen with reference to the opportunity cost of funds, except that for financial analysis, it is the **private** opportunity cost. The discount rate used is the weighted average cost of capital to the entity from whose point of view the financial analysis is being undertaken. The discount rate will include a risk premium required by investors and lenders to compensate them for the risk of loss.

The weighted average cost of capital is calculated by multiplying the tax-adjusted cost of each capital component by its proportional weighting and then summing. The two sources of capital are equity and debt.

The cost of equity capital will be determined by the return demanded by shareholders taking into account their perceptions of the firm's level of risk. The source of the risk premium for shares has been discussed already in terms of systematic risk and the equity premium puzzle.

For borrowed funds, lenders will include a premium for 'default risk' in the interest rate they charge. The default risk premium in the interest rate compensates them for the losses made due to the small proportion of borrowers who will default on their loans. Say a lender has a portfolio of $100m worth of loans and each year, on average, loses one percent of his or her capital ($1.0m) as a result of borrowers defaulting on their loans. The risk-free interest rate is 5 per cent. If the lender charged all customers interest at 6 per cent, the expected rate received would be 5 per cent, the risk-free rate, after taking account of loan defaults.[8] The risk

8 More precisely, the interest rate changed in this situation would be 6.061 per cent, fractionally above 6 per cent (by 0.061 ≈ 0.06/99), because the defaulters fail to pay the interest as well as the principal in the year they default. The formula for the interest rate including default risk is: $(1+r)/(1-\delta) - 1$ where r is the risk-free rate and δ is the probability that a borrower will default.

adjusted for is the downside risk that the borrower will go bankrupt, not the risk of an individual project under-performing.

The costs and revenues in financial analyses should still be central values, free of downside risk.

Crowding out

First, it should be noted that, although using a risk-free discount rate makes public-sector projects appear more desirable, BTRE is, at the same time, advocating measures that work in the opposite direction. Eliminating downside risk **reduces** the relative attractiveness of public-sector projects. So also does use of a certainty equivalent in the rare cases where it is warranted.

If the state-contingent approach to adjusting for risk were used in CBAs instead of risk premiums, it is not clear whether the overall result will be to advantage or disadvantage public-sector projects. If the upward bias of optimistic forecasts were, on average, approximately negated by high discount rates, then the main result of correct treatment of risk should be to alter the ranking of projects. Projects with benefits accruing further in the future would be favoured at the expense of projects with more immediate benefits, because benefits occurring far in the future are disproportionately penalised by a discount rate that includes a risk premium. Such a change should improve the allocation of resources. Admittedly, there is a danger that analysts will use the risk-free discount rate while continuing to make optimistic forecasts. This needs to be guarded against. The next section addresses the broader issue of countering 'optimism bias'.

Second, there is a presumption in the crowding-out argument that the risks and costs associated with a project will be the same regardless of whether it is implemented by the private or the public sector. Take the example of a road project that is part of a publicly owned road network. Decisions made elsewhere in the network will affect the worth of the project in question. If the project were publicly owned, the same organisation would control decisions elsewhere in the network. The level of network risk attaching to the project would therefore be less if it were undertaken by the public sector.

In other cases, network risk considerations will favour private ownership of infrastructure. An example might be a mine operator that is the sole user of a railway line to move minerals from mine to port. If the government provided the railway line, it would need to take account of the risk that the mine will produce below expected output levels. With the private-sector mine operator providing the railway line, investment in the railway line is undertaken with a much better understanding of the risk factors affecting mine output. Furthermore, the mine owner has greater control over the railway line, which is essential to the operation of the mine.

Costs may differ between the public and private sectors as well as risks. The private sector may be able to undertake certain tasks better than the public sector.

Differences in risks and costs can mean that even though the denominators in discounting formulas may favour the public sector due to lack of a risk premium, the numerators may be offsetting. Within the private sector, big business has access to cheaper capital than small business because large firms are able to offer greater security to investors and lenders, but this does not mean that big business invariably crowds out small business. The cost of capital to a firm is one of many factors that determine its overall competitiveness.

Proper accounting for risk in public-sector project evaluation should promote attainment of a division of resources invested by the public and private sectors that maximises the overall economic welfare of society.

A further factor that mitigates against public-sector investment crowding out private-sector investment is the budgetary and political processes that determine levels of government spending, taxing and borrowing. In practice, a positive NPV at some discount rate is not the only test that a project has to pass into order to be implemented. The existence of budget constraints imposes an additional economic hurdle on projects being appraised by CBA. A project normally has to achieve a BCR significantly above unity to be implemented. Hence, the level of the discount rate is by no means the sole constraint on government investment.

Countering ‘optimism bias’

Downside risk in economic evaluation of major infrastructure projects is part of a wider problem called ‘optimism bias’, which has been attracting attention of late (Flyvbjerg et al 2003, HM Treasury 2003, BDOT 2004). Optimism bias extends to the whole suite of techniques used to assess projects — CBA, financial analysis, and assessments of environmental, regional economic, and macro-economic impacts. The causes of optimism bias are more wide-ranging than simple failure to consider what can go wrong. According to Flyvbjerg:

> Theories on cost overrun suggest that optimism bias could be caused by a combination of how the decision-making process is organised and strategic behaviour of actors involved in the planning and decision-making processes. Our analysis indicates that political–institutional factors in the past have created a climate where only a few actors have had a direct interest in avoiding optimism bias. (BDOT 2004, p. 5)[9]

Comprehensive and transparent risk assessment in CBAs using the state-contingent approach, as advocated by this report, should do much to counter optimism bias. BDOT (2004, p. 58) suggests that the risk assessment methodology include:

- generic risk analysis checklists;
- requirements for mandatory risk identification workshops (with multi-disciplinary participation);
- requirements for the use of statistical scenario analysis on large projects; and
- requirements for assessment of the market structure and possible levels of competition.

However, since project evaluation is not costless, it is important to ensure that the amount of effort devoted to risk assessment be commensurate with the size of the project under consideration.

The risk assessment process can still be influenced by the ‘political–institutional factors’ that lead to bias. In any serious attempt to address optimism bias, formal

9 See BDOT 2004, pp. 38–40 for more detail on the causes of optimism bias.

requirements to undertake high-quality risk assessment using the state-contingent approach should be introduced in conjunction with other strategies, such as:

- emphasis on establishing realistic budgeting as an ideal, thereby de-legitimising over-optimistic budgeting as routine (BDOT 2004, p. 57);
- use of fiscal incentives against cost overruns, for example, through shifting the burden of financing project cost escalation to the parties best able to control it (BDOT 2004, p. 57);
- careful and independent review of project appraisals, especially when the appraisals have been carried out by or at the instigation of project proponents; and
- establishment of a post-completion evaluation framework that provides for forecasts of costs, benefits and non-monetary quantities made during the appraisal, planning and design stages to be compared with actual outcomes.[10]

10 ATC 2004 volumes 2 and 3, section 2.17 discusses such a post-completion evaluation framework.

Conclusion

Two broad types of risk have been identified: downside and pure risk. Downside risk comes about from taking a no-surprises approach to forecasting costs and benefits. The resultant estimates tend to be biased in favour of the project proceeding. Pure risk is variation around unbiased estimates of costs and benefits. For both types of risk, adding a risk premium onto the discount rate is generally a poor way to allow for risk in CBA's of public sector projects.

The best approach to minimise downside risk is to use the state contingent approach to ensure that the estimates of costs and benefits are expected values.

Pure risk is likely to be quantitatively of little significance in most cases and so can be safely ignored. The main reasons are that:

- the costs and benefits of public-sector projects tend to be spread over large numbers of people, so poor performance by any one project is unlikely to have much effect on the welfare of any individual;
- since the welfare of each individual is affected by large numbers of projects, poor performance by some is likely to be offset by better-than-expected performance by others; and
- to some extent, project performance is likely to be correlated with the general level of economic activity, an effect that cannot be diluted by projects offsetting one another. Even so, when the size of the variations in economic activity over time are compared with total consumption, the variation is small in relative terms. Similarly, when the size of **variations** in project net benefits arising from changes in the level of general economic activity are compared with total project benefits, again the variation is small in relative terms. The warranted downward adjustment to project benefits therefore turns out to be minuscule.

The exception is where over- or under-performance by a project has a large effect on the welfare of some individuals. In this case, an adjustment can be made using the state-contingent approach combined with the expected utility model.

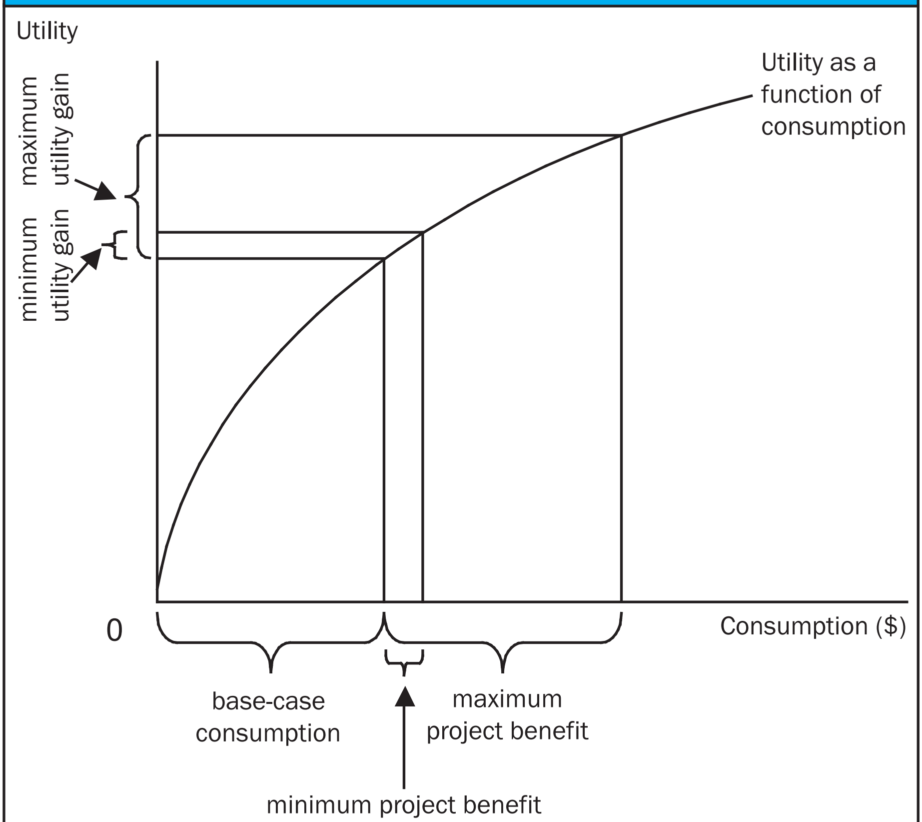
This report has advanced the case made in BTE 1999 for discounting at the real long-term government bond rate without the addition of any risk premium. It is important to understand that this is not a call to ignore risk. Action needs to be taken to ensure that the results of CBAs are **expected** values. Analysts also need to be alert for the rare situations where estimation of a certainty equivalent is justified.

Annex 1

Diagramatic exposition of certainty equivalent and risk premium

The curve in figure 1.1 shows utility as a function of consumption measured in dollars. The function is curved because of 'diminishing marginal utility' — as a person's level of consumption increases, additional dollars provide successively less additional utility. A public-sector investment project is forecast to benefit a given individual by either a large amount or a small amount, depending on risky circumstances. In the base case (without the project), the individual in question will consume at the level shown in figure 1.1.

FIGURE 1.1 CONSUMPTION–UTILITY CONVERSION WITH CERTAIN BASE-CASE CONSUMPTION



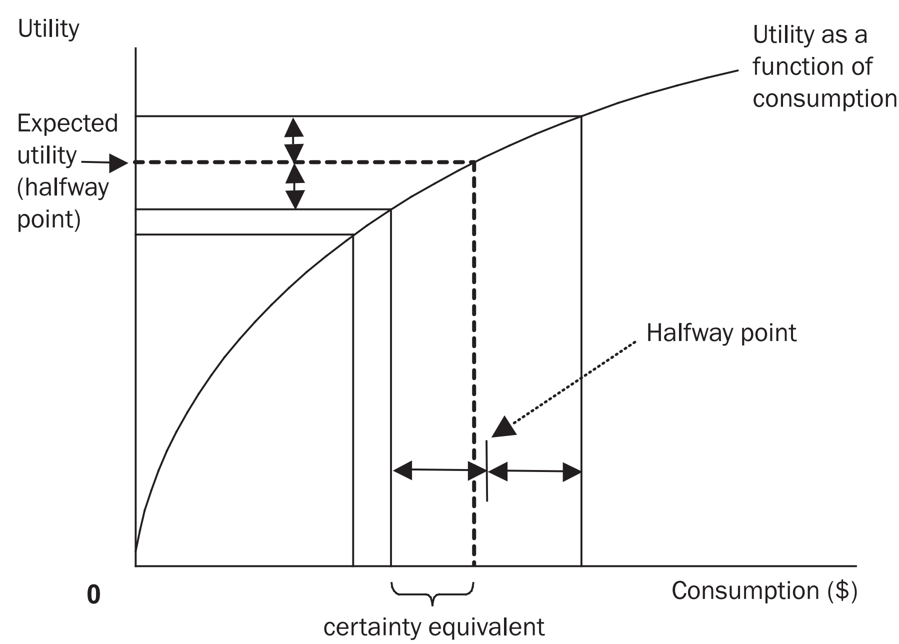
The project will increase his or her consumption by either the minimum or maximum amounts shown. The consumption levels measured in dollars on the horizontal axis can be converted into utility units on the vertical axis using the curve.

Figure 1.2 shows the same curve and consumption levels as for figure 1.1. If there is a 50 per cent probability that the project will achieve the maximum benefit, and a 50 per cent probability that it will achieve the minimum benefit, then the expected utility will lie exactly halfway between the maximum and minimum values. The level of utility represented by the dashed line in figure 1.2 is the expected value. The certainty equivalent level of benefit from the project is found by converting expected utility back into dollars of consumption via the curve, as shown by the dashed lines.

For comparison, the halfway point between the maximum and minimum project-case consumption levels is shown in figure 1.2. It represents the expected value of project-case consumption. As a consequence of the curvature of the utility function, the certainty equivalent is less than the expected value of project-case consumption. Had the utility function been linear, the certainty equivalent would have coincided exactly with the halfway point.

The gap between the certainty equivalent and the halfway point is the risk premium, that is, the amount by which the expected value of project benefits needs to be reduced to adjust for the disbenefit of uncertainty.

FIGURE 1.2 CERTAINTY EQUIVALENT WITH CERTAIN BASE-CASE CONSUMPTION



In figures 1.1 and 1.2, the level of project benefits, at least in the maximum-benefit case, was quite large in relation to base-case consumption. The risk premium arises from the curvature of the utility function over the range of the potential benefits of the project. Had the level of potential benefits been small, the effect of curvature of the utility function on the size of the certainty equivalent would have been small. In the case of the formula for the certainty equivalent derived in annex 2, it is assumed that the potential increases in consumption from the project are sufficiently small in relation to total consumption that the utility function between the base-case and project-case consumption levels can be treated as being linear. The risk premium arises from a different cause — correlation between project benefits and overall consumption. Figures 1.3 and 1.4 show how this occurs.

In figure 1.3, base-case consumption may be large or small depending on risky circumstances. Project benefits have to be high when consumption is high, and low when consumption is low, for the two variables to be positively correlated. To keep the diagram simple, it has been assumed that project benefits are zero when base-case consumption is at its minimum level, and that project benefits are at maximum when base-case consumption is at maximum.

FIGURE 1.3 CONSUMPTION–UTILITY CONVERSION WITH UNCERTAIN BASE-CASE CONSUMPTION

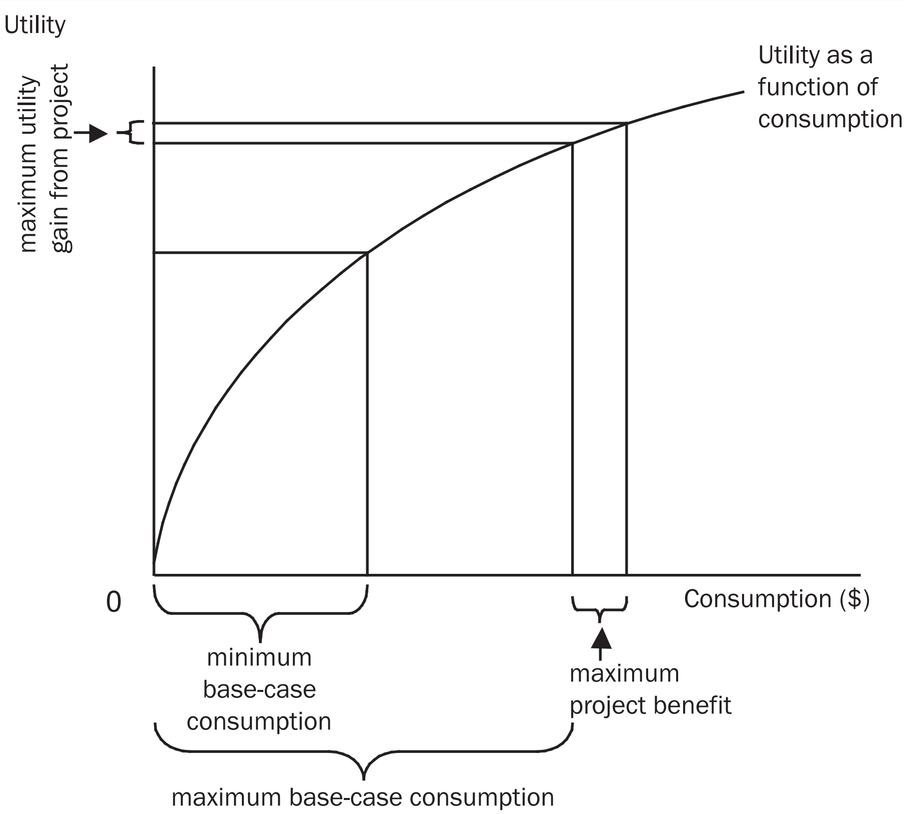
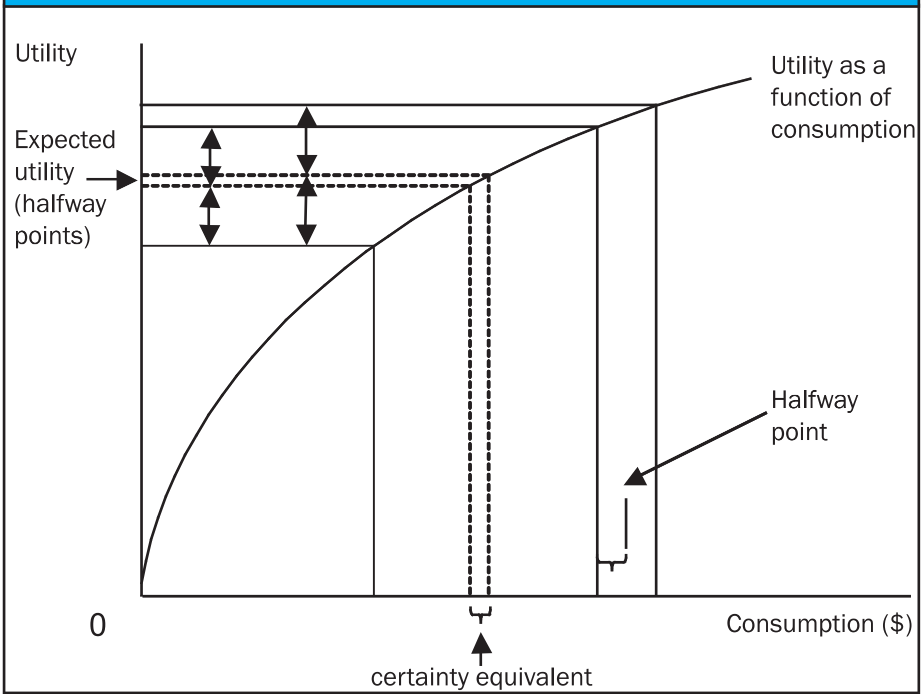


Figure 1.4 is the same as figure 1.3, except that expected utility levels and the certainty equivalent value of the project are shown. It is assumed that there are two states of nature each having a 50:50 chance of occurring. In one state of nature, base-case consumption is at the minimum level and the project, if implemented, fails to yield any benefits at all. In the alternative state of nature, the base-case consumption and project benefits are both at their maximum levels. Hence, there is a perfect positive correlation between base-case consumption and project benefits.

The expected utility level under each scenario can found by taking the halfway point between the maximum and minimum utility levels as shown in figure 1.4. For the base-case, the halfway point is taken between the minimum and maximum base-case utility levels. For the project case, the halfway point is taken between the minimum base-case level (identical with the minimum project-case level) and the maximum project-case utility level. Converting the base- and project-case expected utility levels back into dollars via the utility function, the certainty equivalent value of the benefit is found by taking the difference between the two dollar values.

FIGURE 1.4 CERTAINTY EQUIVALENT WITH UNCERTAIN BASE-CASE CONSUMPTION



For the sake of comparison, a halfway point has been inserted to show the size of the expected value of project benefits without any adjustment for risk. It can been seen that the certainty equivalent is smaller than this amount. The risk premium is the difference between the certainty equivalent and half the maximum project benefit.

This annex has shown diagrammatically two different situations in which a risk premium can arise for economic evaluations public-sector projects:

- uncertain project benefits when the project has potentially a large effect on the welfare of individuals; and
- uncertain consumption levels combined with uncertain project benefits, when project benefits are correlated with consumption.

In practice, it is possible for both situations to occur together.

Annex 2

Mathematical derivation of certainty equivalent and risk premium

In mathematical terms, the certainty equivalent, *e*, of an expected sum of money, *y*, is found by solving the equation:

$$E[u(c_0+e)] = E[u(c_0+y)]$$

where:

- *u(c)* is utility as a function of consumption in dollars; and
- c_0 is the level of consumption in the absence of receipt of *y*.

The critical difference between the two sides of the equation is that *e* is a constant and *y* is a random variable (that it, it has a non-zero variance).

If the utility function was linear, *u(c)= kc* where *k* is a constant, then it can easily be shown that *e* = *E(y)*.

Using a first-order Taylor approximation around c_0 and rearranging:

$$e = E[y] + \frac{\text{cov}\,[u'(c_0), y]}{E[u'(c_0)]}$$

where *u´(c)* is the marginal utility of consumption $\frac{\partial u}{\partial c_0}$.

Note that by making a first-order Taylor approximation, we are assuming that the situation described in figures 1.3 and 1.4 of annex 1 applies, that is, the values of that *y* can take on are small in relation to total consumption. If values of *y* were large in relation to consumption, linear approximation around the locality of c_0 could not be justified.

The expression just derived can be simplified by assuming that an individual's utility function has the form:

$$u(c) = \frac{kc^{1-\gamma}}{1-\gamma} \text{ for } \gamma \neq 1 \text{ or } u(c) = \ln(c) \text{ for } \gamma = 1$$

where k and γ are constants. γ governs the curvature of the function and hence the rate at which the marginal utility of money falls as consumption rises. A value of zero would imply constant marginal utility of money. With a utility function in this form, and making a further approximation:

$$e = \mathrm{E}[y] - \gamma \frac{\mathrm{cov}[c_0, y]}{\mathrm{E}[c_0]}$$

This expression is useful for understanding the factors affecting the certainty equivalent. First, if the marginal utility of money were constant, $\gamma = 0$, the certainty equivalent would equal the expected value. Diminishing marginal utility of money is requisite for risk adversity. Second, with y assumed to be a small proportion of consumption, marginal utility does not change significantly over the range c_0 to $c_0 + y$. The factor causing the certainty equivalent to fall below E[y] is c_0 being a random variable and the marginal utility of income varying inversely with c_0. Receipt of y will be worth more to an individual if c_0 is low than if it is high. Hence, the greater the covariance between y with c_0 the more the certainty equivalent lies below E[y]. Note that in the unlikely event that y was negatively correlated with c_0, the certainty equivalent would exceed E[y].

The risk premium can be defined as the amount a person is willing to pay to avoid a risky situation. From the formula for the certainty equivalent, the risk premium is:

$\mathrm{E}[y] - e = \gamma \frac{\mathrm{cov}[c_0, y]}{\mathrm{E}[c_0]}$, or as a proportion: $\frac{\mathrm{E}[y] - e}{\mathrm{E}[y]} = \gamma \frac{\mathrm{cov}[c_0, y]}{\mathrm{E}[c_0]\mathrm{E}[y]}$.

Annex 3

Estimation of size of risk premium for cost-benefit analyses

Rewriting the expression derived in annex 2 for the risk premium in the notation of the CCAPM, the risk premium expressed as a proportion is:

$$\frac{E[y_s]-e}{E[y_s]}=\gamma\frac{\mathrm{cov}[(c_s),Y_s]}{E[c_s]E[Y_s]}$$

where: c_s and Y_s are, respectively, consumption and the amount to be received in state s, and γ is a constant in the assumed utility function called the coefficient of relative risk aversion. The expectation is taken across all states.

The correlation coefficient between c_s and Y_s is, by definition, $r(c_S,Y_S)=\frac{\mathrm{cov}[c_S,Y_S]}{\sigma[c_S]\sigma[Y_S]}$.

Substituting $\mathrm{cov}[c_s,Y_s]=r(c_s,Y_s)\sigma[c_s]\sigma[Y_s]$ and coefficients of variation $\mathrm{cv}[c_S]=\frac{\sigma[c_S]}{E[c_S]}$ and $\mathrm{cv}[Y_S]=\frac{\sigma[Y_S]}{E[Y_S]}$, the risk premium as a proportion of $E[Y_s]$ is:

$$\frac{E[Y_S]-e}{E[Y_S]}=\gamma\frac{\mathrm{cov}[c_S,Y_S]}{E[c_S]E[Y_S]}=\gamma(c_S,Y_S)\ \mathrm{cv}[c_S]\ \mathrm{cv}[Y_S]$$

The relative standard deviation of aggregate consumption averaged over time is around 3 per cent, that is $\mathrm{cv}[c_s]=0.03$. It might reasonably be assumed, then, that the systematic component of the coefficient of variation of the benefits of a typical project (that is, the component correlated with consumption) is also around 3 per cent, so that $r(c_s, Y_s)\mathrm{cv}[Y_s]=0.03$. Most studies suggest that $0 \le \gamma \le 2$. Say $\gamma = 1$, a central value, then the previous formula for risk premium yields:

$$\frac{E[Y_S]-e}{E[Y_S]}=\gamma(c_S,Y_S)\ cv[c_S]\ cv[Y_S] = 1 \times 0.03 \times 0.03 = 0.0009,$$

a risk premium of approximately 0.1 per cent expressed as a proportion of the expected flow $E[Y_S]$.

Annex 4

First-year rate-of-return criterion for optimum timing of projects[11]

Figure 5.1 shows the basis of the first-year rate-of-return (FYRR) criterion. The block heights represent net benefits in each year during which a project is in place. Benefits grow over time due to growth in demand. The dashed line is set at the discount rate multiplied by the capital cost (rK). The optimal implement time occurs in the first year when $B_1=rK$. For years prior to the optimal time, when $B_1=rK$, delaying the project by one year results in a loss to society of the benefits for that year. However, the capital costs of the project can be invested elsewhere (or not borrowed from overseas saving on interest) saving society the time value of the funds for one year, rK. Society is better off by $rK–B_1$. So society, as a whole, gains from delay of the project. For years after the optimal time, when $B_1>rK$, the value of forgone benefits from delay of the project by one year exceeds the time value of the capital costs. Society is worse off by $B_1–rK$. So delay of the project imposes net costs on society.

If a project passes the FYRR test, $B_1>rK$, its optimal implementation time lies in the past. If it fails, that is, $B_1<rK$, its optimal implementation time is in the future and deferral of the project should be considered.

The FYRR criterion is based on an assumption that benefits do not decline in any years. In figure 5.2, due to a dip in project benefits, there are two possible optimal times. The only way to determine the true optimal time is to treat the two optimal times as mutually exclusive project options, that is, to compare the NPVs with the project implemented at each of the two times when $B_1=rK$.

Another important assumption underlying the FYRR criterion is that changing the year of implementation does not alter project benefits in any years. For example, regardless of whether the project is completed in 2010 or 2015, net benefits in 2020 will be the same. Delaying project construction might also delay times of major maintenance expenses. However, these would be incurred well into the

11 The diagrammatic explanation of the optimal timing condition is reproduced from ATC 2004, volume 3, section 2.12.7.

FIGURE 4.1 OPTIMAL PROJECT IMPLEMENTATION TIME

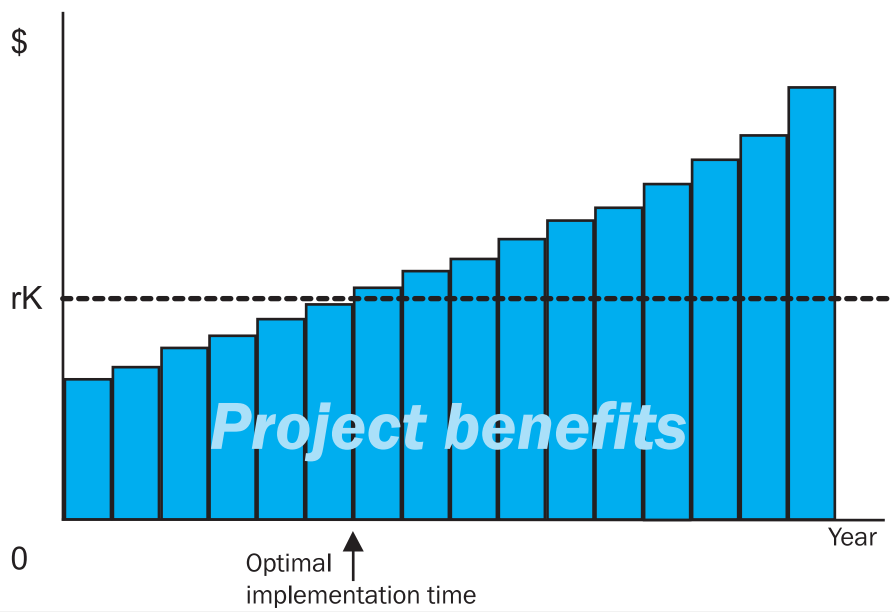
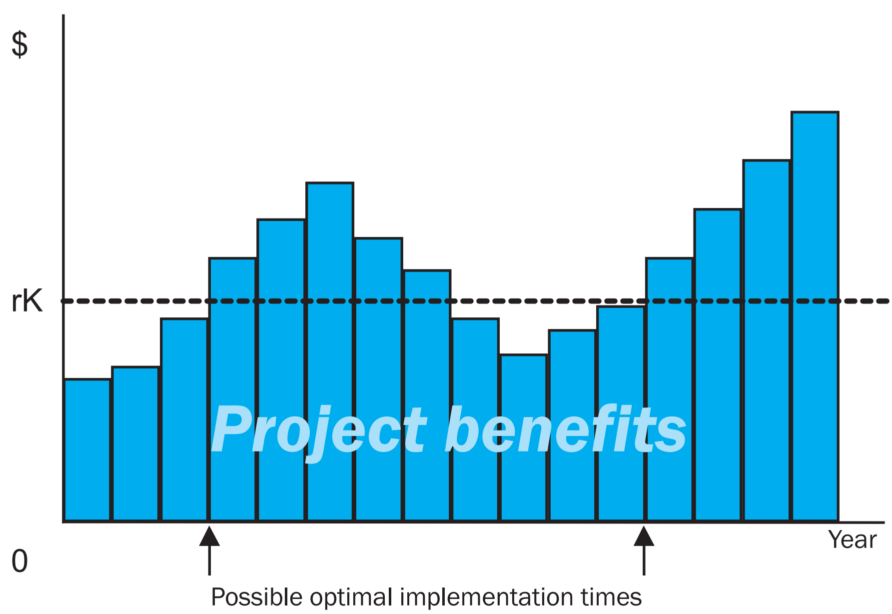


FIGURE 4.2 POSSIBLE OPTIMAL PROJECT IMPLEMENTATION TIMES



future so that the effects of changing their times on a project's NPV would be minimal. Small violations of this assumption are acceptable.

If construction occurs over more than one year, investment costs need to be discounted forward to the year of completion of construction.

Where maintenance costs occur at fixed numbers of years after the implementation time, that is, delay of the implementation time by say five years delays all maintenance costs by five years, the present value of these maintenance costs should be added on to the project capital cost for the purposes of calculating the FYRR. The reason is that the delay of the project leads to additional gains to society from delay of the future maintenance costs. The size of the gain is given by the discount rate multiplied by the present value of these future maintenance costs. However, there may be off-setting additional maintenance costs in the base case if the project is delayed.

Where the effects of delaying the implementation time are more pronounced or more complex, the NPV should be recalculated for each possible optimal implementation year to find the one with the maximum NPV for a given year of analysis. The first-year rate-of-return criterion can be employed to find the possible optimal implementation years.

The mathematical derivation below relies on an assumption that the project has an infinite life. When calculating NPVs with different project implementation times to find the optimal time, it is necessary to discount over a finite time period. As well as cutting out years with low benefits in the near future, delay of the project involves adding in years with high benefits in the far future. The latter changes can distort the choice of optimal time. The remedy is to extend the discounting period sufficiently far into the future so that changing the final year of discounting, as the project is delayed, has negligible effect on the net present value. The lower the discount rate is in relation to the annual growth rate for benefits, the longer the required time period. In extending the discounting period, it may be necessary to add in investment costs in the future to allow for replacement of ageing infrastructure. Delay of project implementation will delay these replacement costs, which is a benefit from project deferment that ought to be included in the analysis.

Mathematical derivation

It is assumed that benefits from the project as a function of time, *B*(*t*), are continuous and upward sloping, and that the project has an infinite life. With investment occurring at time t^* and continuous compounding, the net present value of costs and benefits is:

$$NPV = \left[\int_{t^*}^{\infty} B(t)e^{-rt}dt\right] - Ke^{-rt^*}.$$

where *K* is the capital cost and *r* is the discount rate. If this equation is differentiated with respect to t^*, the result is:

$$\frac{dNPV}{dt^*} = -B(t^*)e^{-rt^*} + rKe^{-rt^*}.$$

The optimum time to invest is found by setting this equal to zero, which leads to the result:

$$B(t^*) = rK \text{ or } \frac{B(t^*)}{K} = r.$$

The second order condition for a maximum is that, in the region of the optimum:

$$-e^{-rt}\frac{dB}{dt} < 0, \text{ which holds if } \frac{dB}{dt} > 0.$$

Hence, the FYRR criterion depends on an assumption that project benefits increase over time in the region of the optimum.

Annex 5

Example of estimation of a certainty equivalent

A project to provide improved access to a farming district will significantly increase farm incomes. However, if there is serious flooding, farm incomes will be lower due to crop losses and the improved road, which provides only limited flood immunity, will yield no benefits at all. The probability of serious flooding is 0.2. The farms are grouped according to income levels. For one group, the average consumption levels are as follows:

	($'000 per annum)	
No flooding	$100	$150
Flooding	$60	$60

As the relative changes in consumption levels are substantial, the expected utility approach is warranted.

Table 5.1 shows how the certainty equivalent, the 'total expected dollars' is estimated for five values of γ: 0.0, 0.5, 1.0, 1.5 and 2.0. The formulas used to derive the table are provided below.

Consumption levels for each scenario for the base and consumption cases are first converted into utility units. With c_0 set at $100,000 in each value of γ, consumption is valued as 100 utility units when the dollar value of consumption is $100,000. The scaling factor, k, is varied for each value of γ to ensure this.[12] Expected utility is estimated for the base case and for the project case. These expected utilities are converted back into dollars, which are the certainty equivalents. Finally, the difference between the certainty equivalents in the base case and project case is taken to obtain the certainty equivalent of the benefit.

12 All monetary values in table 5.1 are thousands of dollars. Had the calculations been undertaken in dollars instead, the values of k would have been different.

It is essential that the conversion from utility units to back into dollars be undertaken **before** taking the difference between the base and project cases, not **after**. Otherwise, a wrong result will ensue because the conversion from utility to dollars will have been made using a part of the utility function that applies only at low levels of income.

The case where γ = 0.0 shows the risk neutral situation. The utility function is a straight line and estimation of certainty equivalents makes no difference to the final result. As levels of γ are increased above zero, the utility function becomes more curved, risk aversion increases, and the downward adjustment to benefits becomes greater.

For values of γ above one, the number of utility units becomes a decreasing function of consumption. This does not contradict the normal assumption of economics that more is preferred to less. It means that under this particular way of measuring utility, smaller numbers of utility units indicate higher levels of satisfaction and conversely. The estimates of benefits for γ = 1.5 and γ = 2.0 have the expected signs and magnitudes.

Formulas for table 5.1

$k = (1-\gamma)c_0^{\gamma}$ for $\gamma \neq 1$ and $k = \dfrac{c_0}{\ln(c_0)}$ for $\gamma \neq 1$.

c_0 = $100,000 for all values of γ

utility = $\dfrac{kc^{1-\gamma}}{1-\gamma}$ for $\gamma \neq 1$ and utility = $k\ln(c)$ for $\gamma = 1$

expected utility = $\sum_{s=1}^{n}$ probability$_s$ x utility$_s$ for all states of nature s from 1 to n, where $\sum_{s=1}^{n}$ probability$_s$ = 1.0

consumption in dollars $= \left[\dfrac{(1-\gamma) \text{ x utility}}{k}\right]^{\frac{1}{(1-\gamma)}}$ for $\gamma \neq 1$ and

consumption in dollars $= \exp\left[\dfrac{\text{utility}}{k}\right]$ for $\gamma = 1$

Table 5.1 Example of estimation of a certainty equivalent

γ

k = 1.0		*base case*	*project*	*base case*	*project*	*base case*	*project*
no flooding	0.8	$100	$150	100	150	80	120
flooding	0.2	$60	$60	60	60	12	12
Total expected utility						92	132
Certainty equivalent ($'000)						$92	$132
Benefit (project minus base case) ($'000)							$40

γ

k = 5.0		*base case*	*project*	*base case*	*project*	*base case*	*project*
no flooding	0.8	$100	$150	100	122	80	98
flooding	0.2	$60	$60	77	77	15	15
Total expected utility					95	113	
Certainty equivalent ($'000)						$91	$129
Benefit (project minus base case) ($'000)							$38

γ

k = 21.7		*base case*	*project*	*base case*	*project*	*base case*	*project*
no flooding	0.8	$100	$150	100	109	80	87
flooding	0.2	$60	$60	89	89	18	18
Total expected utility						98	105
Certainty equivalent ($'000)						$90	$125
Benefit (project minus base case) ($'000)							$35

γ

k = – 500.0		*base case*	*project*	*base case*	*project*	*base case*	*project*
no flooding	0.8	$100	$150	100	82	80	65
flooding	0.2	$60	$60	129	129	26	26
Total expected utility						106	91
Certainty equivalent ($'000)						$89	$120
Benefit (project minus base case) ($'000)							$31

γ

k = – 10000.0		*base case*	*project*	*base case*	*project*	*base case*	*project*
no flooding	0.8	$100	$150	100	67	80	53
flooding	0.2	$60	$60	167	167	33	33
Total expected utility					113	87	
Certainty equivalent ($'000)						$88	$115
Benefit (project minus base case) ($'000)							$27

References

Asia Development Bank 2005, *Guidelines for the Economic Analysis of Projects,* ADB, Manilla. Available at <http://www.adb.org/Documents/Guidelines/ Eco_Analysis/default.asp>

Australian Transport Council 2004, *National Guidelines for Transport System Management in Australia,* 3 volumes, Department of Transport and Regional Services, Canberra. Available from the website of the Australian Transport Council, <http://www.atcouncil.gov.au/atcdocs.htm>

Austroads 2002, *Economic Evaluation of Road Investment Proposals: Risk Analysis,* Austroads, Sydney. Available from <www.austroads.com.au>

Bannock, G., Baxter, R.E., and Davis, E. 2003, *Penguin Dictionary of Economics,* 7th edition, London.

British Department of Transport 2004, *Procedures for Dealing with Optimism Bias in Transport Planning: Guidance Document,* Report prepared by B. Flyvbjerg in association with COWI. Available from the website of the British Department of Transport,<http://www.dft.gov.uk/stellent/groups/dft_localtrans/documents/page/dft_localtrans_029645-01.hcsp>

Brealey R.A. and Myers S.C. 2003, *Principles of Corporate Finance,* 7th edition, McGraw Hill, Boston, Mass.

Bureau of Transport Economics 1999, *Facts and Furphies in Benefit–Cost Analysis: Transport,* Report 100, BTE, Canberra.

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

HM Treasury 2003, *Supplementary Green Book Guidance: Optimism Bias.* Available at www.hm-treasury.gov.uk/greenbook.

Leuhrman, T.A. 1998 Investment Opportunities as Real Options: Getting Started on the Numbers, *Harvard Business Review,* July–August, pp. 51–67.

Appendix

Risk and discounting in project evaluation

John Quiggin

Australian Research Council Professorial Fellow, School of Economics
Faculty of Business, Economics and Law, University of Queensland

I thank Peter Forsyth, Mark Harvey, Phil Potterton, Quentin Reynolds and Nancy Wallace for helpful comments and criticism.

Summary

The central object of this paper is to address the treatment of uncertainty in the application of benefit–cost analysis to project evaluation and to assess the appropriateness of the practice of using a higher discount rate in the evaluation of projects that are regarded as more risky.

The problem is approached using a state-contingent framework which clearly displays the symmetry between the problem of evaluating projects under uncertainty and the more familiar problem of discounting flows of costs and benefits received over time. In this framework, the value of additional income flows received in any time period and state of nature depends on the marginal utility of income in that state of nature. Both discounting of future benefits and risk aversion may be explained as a result of the observation that the marginal utility of income is diminishing in the level of income.

A crucial distinction is drawn between downside risk and pure risk. Pure risk refers to variation about a mean value, and is typically measured by the variance of a random variable. In the context of financial evaluation, pure risk is commonly subdivided into systematic risk (risk correlated with variations in aggregate consumption or aggregate returns to capital) and idiosyncratic or diversifiable risk (risk that is not correlated with aggregate consumption and can therefore, in principle, be eliminated by diversification).

The concept of downside risk refers to outcomes that are worse than some median or 'no surprises' estimate. This is the concept of risk that is most commonly used in general discussion of risk. Downside risk frequently arises in project evaluation if insufficient attention has been paid to possible adverse outcomes.

The analysis yields three main conclusions. First, in most project evaluations, downside risk is more significant than pure risk. Second, for public sector projects, the economic cost of pure risk is small. Third, particularly in the case of downside risk, it is inappropriate to deal with risk by applying an adjusted discounted rate to a projection based on median or modal values. The best procedure is to represent uncertainty explicitly in order to generate unbiased estimates of mean value.

Risk and discounting in project evaluation

Introduction

The treatment of uncertainty is one of the most problematic areas in benefit–cost analysis. The most common approach employed in practice, that of using a higher discount rate in the evaluation of projects that are regarded as more risky, has generally been recognised as unsatisfactory (Little and Mirrlees 1974, Sugden and Williams 1985, Department of Finance 1991). However, there is no generally accepted alternative.

The object of this paper is to explore these issues using a formal representation of uncertainty in terms of actions with state-contingent outcomes. This approach clarifies the analogy between the problem of evaluating uncertain income flows and the problem of discounting income flows received over time. Since the discounting problem is more widely-understood, this analogy assists evaluation of uncertain income flows.

Two crucial insights emerge from this analysis. The first is the importance of distinguishing between 'pure risk' and 'downside risk'.

Pure risk is variation about a mean value, commonly measured by the variance. In the context of financial evaluation, pure risk is commonly subdivided into systematic risk (risk correlated with variations in aggregate consumption or aggregate returns to capital) and idiosyncratic or diversifiable risk (risk that is not correlated with aggregate consumption and can therefore, in principle, be eliminated by diversification).

Downside risk reflects the possibility of adverse events that reduce the net returns of a project. To illustrate the distinction, consider the possibility that an individual is given a free ticket in a lottery. In ordinary language, this would not be considered as an increase risk, but in statistical terms, both the mean and variance of wealth have increased. In this example, pure risk has increased, but downside risk has not.

The best way of dealing with downside risk is through the explicit representation of uncertainty in terms of an event tree, taking account, as far as possible, of future contingencies that may affect project returns. Because most surprises are unfavourable, the mean estimates of returns derived from an analysis will generally be lower than those obtained from a 'surprise-free projection', in which variables are assumed to take their modal or most likely values.

Hence, project analysis should be undertaken as far as possible on the basis of mean estimates rather than modal or surprise-free projections. Attempts should be made, to ensure that evaluations of projects competing for funding from the same source should be of equally good quality, particularly with respect to contingency analysis. After these steps are taken, it is likely that some upward bias in estimates of the benefits of projects will remain. The optimal response is to adjust estimates of the benefits of projects based on past experience of the average divergence between anticipated and observed outcomes rather than to adjust discount rates.

The second crucial insight relates to the treatment of pure risk. It is shown that many of the central debates about risk and discounting reflect implicit presumptions about the 'equity premium puzzle', identified by Mehra and Prescott (1985). The core of the equity premium puzzle is the empirical observation that the market premium for investment in risky equity is very large – typically equal to about 50 per cent of expected returns. By contrast, the risk premium derived from a model based on the assumption that individuals rationally optimise their consumption by trading in efficient capital markets is so small as to be negligible in most cases.

It will be argued that most plausible explanations of the equity premium puzzle involve violations of the efficient market hypothesis. Such violations weaken the presumption that prices and rates of return observed in financial markets should form the basis of project evaluation for the public sector. For most purposes, the analysis of public investments should be based on the assumption that risks are spread through the tax system in a way which yields a negligibly small risk premium.

The paper is organised as follows. Section 1 is a review of discounting under certainty, including notions of present and future values and the net present value concept. The implications for project selection are considered with particular emphasis on the costs of displaced consumption and investment in the presence of distorting taxes and other transactions costs. The implication arising from these cases is that, in the presence of transactions costs or other obstacles to intertemporal trade in consumption flows, the distribution of costs and benefits cannot be disregarded. Different individuals may face different prices, and this fact must be taken into account.

Section 2 deals with the nature of risk. Sources of risk in infrastructure projects, including construction costs, operation costs, demand risk and network risk are described. Risk is addressed in a state-contingent framework. The key components of this framework, including states of nature, events and random variables are described and the associated analytical tool of event-tree analysis is illustrated. The distinction between downside risk and 'pure' risk is described. It is argued that downside risk arises naturally when the parameters used in project evaluation are derived from modal estimates or 'no surprises' scenarios. Within the category of pure risk, a distinction is drawn between systematic and idiosyncratic risk, particularly in relation to demand risk. The expected utility model is described, and its appropriateness as a basis for project evaluation is defended.

Section 3 is the analytical core of the paper, showing how the state-contingent approach to project evaluation under uncertainty may be implemented. The crucial steps are identified as: specifying the state space; estimating income flows; and determining state-contingent prices. An example is used to illustrate these steps. A range of practical issues is discussed, including the use of scenarios and sensitivity analysis. Attention is then turned to the common practice of using higher discount rates to 'adjust' estimates of net present value for the presence of risk. It is argued that, while this approach may be a reasonable approximation in some cases, particularly in relation to demand risk, it is not appropriate in general. The use of higher discount rates to offset downside risk arising from the use of mis-specified parameter estimates is shown to be even less defensible. Some practical responses to problems of bias in the parameters used in project evaluation are discussed.

Section 4 deals with the cost of pure risk, and yields the conclusion that, for many projects, this cost is negligible. The section begins with a description of the CAPM and CCAPM models, identifying the 'market risk premium for equity' as the crucial parameter in these models. This observation is followed by a survey of the literature on the 'equity premium puzzle', that is, the fact that rates of return to equity are considerably higher than would be expected if capital markets were perfectly efficient. To the extent that the equity premium arises from market failure, rates of return to private equity are not an appropriate guide to the treatment of risk in the evaluation of public projects. The resulting case for disregarding pure risk in public projects is assessed. Some implications are derived for the design and analysis of projects involving co-operation between the public and private sectors. The role of flexibility and option value is considered.

Finally, some concluding comments are offered.

1 A review of discounting under certainty

This section has two objects. The first is to remind readers of the basic principles of discounting under certainty and to emphasise the point that a discount rate is not an exogenously given reality but a way of summarising relative prices of current and future consumption. The second is to prepare notation and terminology for a parallel treatment of uncertainty.

The treatment may appear to labour some points that are painfully obvious to anyone with a basic training in economics. However, the discussion of benefit–cost analysis under uncertainty has been bedevilled by confusion about such basic issues.

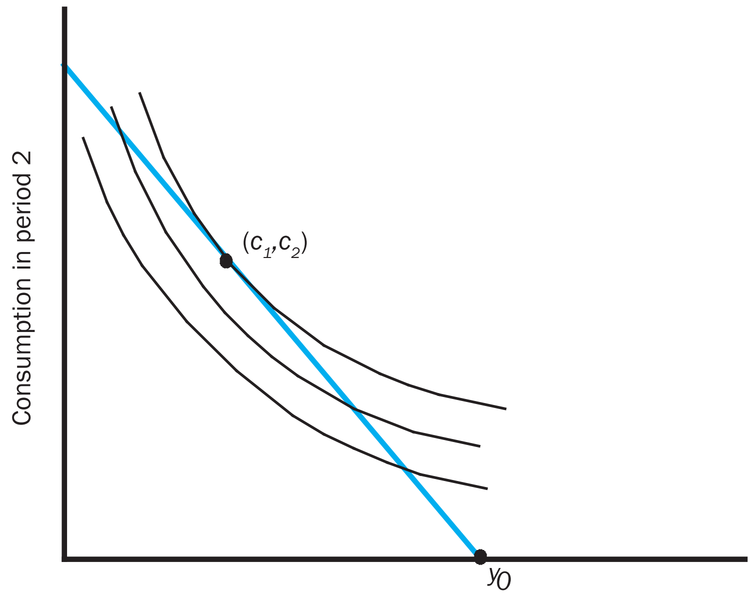
1.1 Present and future values

The simplest possible investment project is one in which an individual makes an investment yielding a definite payoff at a fixed date in the future, say, in one year's time. That is, the individual gives up some consumption now in order to increase consumption in the future. For concreteness, suppose that a kilogram of potatoes may be consumed today or replanted to yield two kilos in a year's time.

The central insight of the neoclassical economic theory of interest, which may be traced back to Fisher (1930) is that this choice is, in essence, the same as a choice between items of present consumption, say, wine and bread. Hence, just as in the case of present consumption, the preferences that determine choice between present and future consumption may be represented in terms of indifference curves.

This point is illustrated in Figure 1. The curved surfaces are indifference curves, that is, combinations of current and future consumption that are regarded by a decision-maker as equally valuable.

FIGURE 1 CONSUMPTION CHOICES OVER TIME



In the absence of trade, the trade-off between current and future consumption will be different for different individuals, depending on their preferences and endowments. The stronger an individual's preference for current over future consumption, and the smaller their initial endowment of present consumption relative to future consumption, the higher will be their marginal rate of time preference.

Consider, on the other hand, the case when all individuals can freely trade present and future consumption (that is, borrow or lend) at a common price, say *p* units of future consumption for each unit of present consumption. The standard convention is to express this price in terms of a rate of interest *r*, such that:

(1) $p = 1+r.$

Another common convention is to use a discount factor

(2) $\beta = \frac{1}{p}.$

The optimal solution for an individual is illustrated in Figure 1. The initial wealth level y_0 determines a budget set. The highest available indifference curve is represented by the point of tangency at (c_1, c_2). At this point the slope of the indifference curve is equal to the discount factor

(3) $\frac{\partial c_1}{\partial c_2} = \beta.$

If the individual's preferences are represented by an objective function $W(c_1, c_2)$, the equilibrium condition is that:

$$(4)\qquad \frac{\partial W}{\partial c_1} = \beta \frac{\partial W}{\partial c_2},$$

or, conversely

$$(5)\qquad p\frac{\partial W}{\partial c_1} = \frac{\partial W}{\partial c_2}.$$

The most common form for the objective function is one characterised by a utility function u, such that

$$(6)\qquad W(c_1, c_2) = u(c_1) + \frac{u(c_2)}{(1+\delta)}$$

$$(7)\qquad \frac{\partial W}{\partial c_1} = u'(c_1);\text{ and}$$

$$(8)\qquad \frac{\partial W}{\partial c_2} = \frac{u'(c_2)}{(1+\delta)}$$

In general, we expect a positive rate of interest. This may arise for one of two reasons. First, given the existence of productive investment opportunities, we expect consumption to be growing over time, and therefore $c_2>c_1$. Assuming diminishing marginal utility of consumption, then $\partial W/\partial c_2 < \partial W/\partial c_1$. That is, since consumption in period 2 is higher than in period 1, the value of a marginal unit of consumption is less.

The second possibility is that $\delta>0$. That is, utility received in the future is considered to be inherently less valuable than utility received now. From a social point of view, this assumption violates requirements for intergenerational equity. Attention in this paper will be focused on the case for discounting arising from diminishing marginal utility.

1.2 The concept of net present value

A typical investment project lasts for more than two periods. The central tool of benefit–cost analysis is the notion of present value (often called net present value in the standard case where an initial reduction in consumption is used to generate a flow of consumption services into the future). The present value approach is based on the derivation of a set of prices for future consumption, expressed in terms of present consumption. In effect, this reduces all flows to a common 'currency'.

We will denote the current period as t=0 and subsequent periods as t=1 ...T where T is a time horizon for analysis. (T may be taken to be infinite without affecting the analysis.) We will refer to the flows associated with the project as 'income', denoted by y. Consumption is given by

(9) $c = \underline{c} + y$,

where $\underline{c}$ is consumption in the absence of the project and y is the flow of net income from the project.

Income flows are evaluated by the present value formula

(10) $V = \Sigma p_t \, y_t$,

where p_t is the price of consumption t periods from the present;

the price of current consumption is p_0=1; and

y_t is the net flow of income from the project in period t.

The simplest case of the present value method arises when the interest rate is constant. In a market equilibrium this will be true if consumers have a common, constant, intertemporal elasticity of substitution, and income growth is determined by a constant exogenous rate of technological progress. In practice, there is considerable variation in market interest rates. However, real interest rates rarely stay above 5 per cent or below zero for any extended period, and usually remain in a range between 2 and 4 per cent.

If there exists a constant real interest rate, expressed in annual terms, the relative price of consumption in different years depends simply on the number of years between them (denote this gap by t). A unit of consumption in year τ has a value of $(1+r)^k$ units of consumption in year $\tau+k$. Hence, the present value formula may be written as

(11) $V_\tau = \Sigma_t \, (1+r)^{-t} y_{t+\tau}$

where V_τ is the present value at time τ. If, as is usually the case, τ=0, we can simply write

(12) $V = \Sigma_t \, (1+r)^{-t} y_t$

$=\Sigma_t \, \beta_t \, y_t$

where $\beta_t = (1+r)^{-t}$ is the discount factor for time t.

In the case of constant real interest rates, the evaluation of a project in terms of positive or negative present value will be the same whenever the project is commenced, assuming annual benefits and costs are independent of the starting date. Although real interest rates vary over time, such variations are hard to predict in advance. Hence, most project evaluation is based on the assumption of a constant real rate of discount equal to the current real rate of interest.

Example 1: A simple example may prove useful. We consider a road project with a life of ten years. In the first two years construction is undertaken, and in the following eight years, the road provides transport services by reducing the time taken to travel between the points it connects. (More details on the assessment of such benefits are given by Bureau of Transport Economics 1999). We will assume that all benefits are expressed in real terms and that the real rate of interest is 5 per cent. The second and third columns of Table 1 give the data on the project, expressed in terms of annual flows of costs and benefits for the ten-year project life. The fourth column shows the flow of net returns. The fifth column shows the discount factor in year *t*, calculated as $(1+r)^{-t}$, , where r = 0.05. The final three columns show the discounted cash flows. The final row of the table shows totals for the life of the project.

As illustrated in Table 1, the project has a positive net present value given the discount rate of 5 per cent. Other assessment criteria may also be calculated. The project's benefit–cost ratio is 752/681 = 1.1. The internal rate of return, discussed below, is 7.5 per cent.

Table 1 Example of project evaluation

0	300	0	-300	1.00	300	0	-300
1	400	100	-300	0.95	381	95	-286
2		100	100	0.91	0	91	91
3		102	102	0.86	0	88	88
4		104	104	0.82	0	86	86
5		106	106	0.78	0	83	83
6		108	108	0.75	0	81	81
7		110	110	0.71	0	78	78
8		113	113	0.68	0	76	76
9		115	115	0.64	0	74	74
Total	700	958	258		681	752	71

a Present Value

b Net Present Value

1.3 Implications for project selection

The analysis above is based on the assumption that the individual or group undertaking the project or receiving its benefits can trade current and future consumption freely at fixed prices p_t, that is, borrow or lend as much as they wish at given rates of interest. Moreover, it has been implicitly assumed that all projects can be evaluated separately. In these circumstances, the optimal investment rule generated by the net present value concept is simple: projects should be undertaken if, and only if, they have a positive (net) present value.

In some cases, projects cannot be evaluated separately. One project may be an alternative to another. Alternatively, one project may be feasible only if one or more other projects are undertaken previously or concurrently. The general rule derived from the present value method is that all feasible combinations of projects (sometimes called investment portfolios) should be assessed, and the option yielding the highest (net) present value should be selected.

A related case is that where there is a constraint on the total amount that can be invested in period 0. In this case, it is appropriate to employ a 'shadow price' to take account of the budget constraint.

1.4 Costs of displaced consumption and investment

The assumption that all individuals trade current and future consumption freely at fixed prices (given by p_t for consumption in period t) will not be satisfied if such trades are subject to transaction costs and taxes. The first such case to be discussed in the literature on benefit–cost analysis was that where taxes are imposed on company profits and income from savings. As a result, the pretax return on marginal private investment projects will be greater than the post-tax return to private savings. Marglin (1963, 1967) refers to the post-tax return to private savings as the social rate of time-preference. Similarly, the pretax return on marginal private investment projects may be referred to as the social rate of opportunity cost.

Under these circumstances, consider two possible projects. The first is financed entirely by displacing marginal private investments. For this project, it seems clear that the appropriate discount rate for the calculation of the present value is the pretax return on marginal private investment projects. (Note that problems of risk and uncertainty may modify this presumption, as discussed below.) If on the other hand, a project is financed entirely by reductions in present consumption, that is, by increases in savings, the social rate of time preference is appropriate. Most projects involve a mixture of sources of financing.

A number of possible approaches to the analysis of discounting in the presence of distortions have been proposed (Marglin 1963; Feldstein 1972). When correctly applied, all of these approaches are logically equivalent and yield the same recommendations. The 'shadow price' approach proposed by Marglin clarifies the relationship between project appraisal and the ultimate objective of maximising welfare.

A crucial feature of Marglin's approach is the need to consider both the source of funds used to pay the costs of a project and the way in which the benefits of a project are disposed, for example, whether they are consumed or reinvested.

Marglin proposes the use of shadow prices to express the costs of forgone investment in terms of forgone streams of consumption.

Using the shadow pricing approach, and assumptions of the disposition of project benefits, the impacts of any project can be represented as a sequence of changes in consumption. The present value can then be computed using the social rate of time preference as a discount rate.

A second case of interest arises if the costs of a project are paid by governments but benefits are consumed directly, with no user charges. In this case, it is necessary to take account of the costs of raising revenue, including administration and compliance costs and the deadweight costs of distorting taxes. As in the case of taxes on savings and profits, the theoretically appropriate remedy is to use a shadow price to express all costs and benefits in terms of consumption.

Campbell (1997) suggests that the average deadweight loss associated with taxation revenue is equal to 24 per cent of the revenue raised. (See also Campbell and Bond 1997.) This implies that a project financed entirely from taxation revenue, and yielding benefits entirely in the form of higher consumption, should be approved if and only if the benefit–cost ratio (see section 1.5) exceeds 1.24. More commonly, costs and benefits of projects will include a mixture of consumption and government revenue. Shadow prices can be used to express all flows in terms of consumption streams.

A third possibility arises in an open economy where governments with good credit ratings can borrow freely at a rate of interest determined in international capital markets. This rate of interest then constitutes the opportunity cost of capital.

Finally, it is useful to consider the case when individuals are either constrained from borrowing or face rates of interest higher than the government bond rate r. In this case, the value of consumption benefits provided to individuals will depend on the timing of those benefits. To the extent that public projects lead to smoother consumption streams, their benefits will be greater.

The crucial implication arising from these cases is that, in the presence of transactions costs or other obstacles to intertemporal trade in consumption flows, the distribution of costs and benefits cannot be disregarded. Different individuals may face different prices, and this fact must be taken into account.

1.5 Alternatives to the present value method

The present value method is not the only approach used in the evaluation of investments. Among the alternatives are the payback period, the internal rate of return (IRR), and benefit–cost ratios. Several of these approaches were in use before the development of the present value method and yield results consistent with those of the present value method in certain special cases. In this section, it will be argued that where these alternative methods do not agree with the present value method (assuming appropriate prices), the answer given by the present value method is to be preferred.

The most commonly used alternative to the present value method is based on the IRR. In mathematical terms, the internal rate of rate is simply the dual of the present value approach, assuming a fixed annual rate of return. Rather than determining whether the present value is positive or negative at a given rate of return, the IRR approach determines the rate of interest for which the present value is exactly zero. The associated evaluation procedure is to accept all those projects with an IRR in excess of a 'hurdle rate'.

In the most common cases, the IRR approach yields the same recommendations as the present value approach. Suppose the project involves an initial outlay, so $y_0<0$, followed by a stream of positive returns $y_t \geq 0$. The higher the rate of interest r, the lower will be the present value. In particular, there is a unique rate of interest (the IRR) for which the present value is zero.

Now compare a present value analysis using a discount rate r and an IRR analysis using the same r as the hurdle rate. The project will have a positive present value if and only if the IRR is greater than the hurdle rate. Thus, the two criteria will agree.

Suppose on the other hand, the project is something like a quarry, with a flow of positive returns while extraction takes place, followed by a cleanup cost at the end of the project. In this case, higher discount rates will yield higher present values. Applying the IRR approach to projects of this kind yields exactly the wrong answer. The worse a project is in present value terms, the higher the IRR at which it breaks even.

A more general difficulty with the IRR is that it provides little useful guidance in choosing between alternative, but incompatible, projects, such as alternative routes for a road. If the lifetime of the projects, or even the pattern of benefits over time, is different, there is no reason to suppose that the project with a higher IRR will yield greater benefits in terms of net present value.

One useful feature of the IIR is that its calculation depends only on the flows of the project – the discount rate need not be known in advance. This means that, for projects where the IRR and present value methods agree, the IRR summarises the most relevant information in a single statistic. Hence, calculation of the IRR may be a useful practical tool.

An additional useful characteristic of the IRR is the fact that it gives a rough idea of the robustness of the conclusions of a present value analysis. If the discount rate is 5 per cent and a project has an IRR of 15 per cent, it seems safe to conclude that variations in the parameters used to evaluate the project are unlikely to reverse a favourable evaluation. By contrast, without separate information on the scale of the project, the information that it has a net present value of $10 million gives little insight into the robustness of the claim that the net present value is positive.

These convenient features of the IRR should not be taken to support the view that the IRR is an alternative to present value, with advantages and disadvantages. When the IRR is consistent with a present value analysis, it is a convenient summary statistic for that analysis. When the IRR approach yields conclusions that differ from those of present value analysis, the answer derived from IRR is incorrect. Provided appropriate prices are used, the present value method is always the optimal discounting procedure.

Many of the advantages associated with the IRR may be obtained, with less danger of error, using the benefit–cost ratio, that is the ratio of the present value of project benefits to the present value of project costs. The benefit–cost ratio is greater than one if and only if the present value of benefits exceeds the present value of costs, that is, if and only if net present value is positive.

The benefit–cost ratio is often used for ranking projects in the presence of a budget constraint, in which case the critical ratio at which projects are implemented may be greater than one. If all costs are subject to the budget constraint, but benefits flow to consumers, the critical value may be interpreted as the shadow price of public expenditure. On the other hand, if all flows are financial, operating costs must be regarded as an offset to benefits, and left in the numerator for calculations of the benefit–cost ratio, which might more properly be referred to, in this case, as the ratio of net benefits to construction costs. In this case, the critical ratio must be interpreted as the shadow price of funds in the construction period.

2 The nature of risk

2.1 Introduction

The problems of discounting benefits over time are now relatively well-understood. By contrast, the treatment of issues involving risk remains difficult and controversial. In fact, however, the two problems are very similar. In this section, a state-contingent approach to the analysis of project evaluation under uncertainty will be developed, with particular reference to the risks associated with infrastructure projects.

2.2 Sources of risk in infrastructure projects

To provide a concrete basis for discussion it is useful to focus on the case of a transport infrastructure project, such as a road or rail line. Such a project must be assessed as part of a broader transport network. However, in most cases, it is possible to identify substantial costs and benefits specifically attributable to the project. In most cases, projects of this kind have the 'standard' structure in which an initial period of net outlays (the construction phase) is followed by a period of positive net returns in which the value of services flowing from the project exceeds the costs of operation and maintenance. Given this structure, it is possible to identify separately a range of sources of risk.

2.2.1 Construction costs

Proposals to undertake a transport infrastructure project typically include an estimate of the costs of construction. However, this estimate may turn out to be an underestimate because of increases in input costs, or because of unforeseen technical difficulties, such as equipment breakdowns and adverse weather. In an economic sense, failure to complete the project on time reduces the present value of the services provided by the project and therefore increases the effective cost of the construction phase. Less frequently, things may turn out better than expected, with the project being completed 'on time and under budget'.

2.2.2 Operation costs

After completion of the construction phase, an infrastructure asset must be maintained. In addition, the operator may provide a range of operational services using the asset. For some assets, such as roads, costs of operation and maintenance are relatively stable and predictable and are small relative to initial costs of construction. For other assets, such as airports, operations may be complex and subject to substantial risk.

Another important issue regarding risk and operational costs is the relationship between the construction and operation phases. In some cases, decisions made in the construction phase, for example regarding the quality of materials, may have a substantial impact on subsequent costs of operation and maintenance. In such cases, contractual arrangements in which the constructor is required to undertake maintenance may be appropriate. In other cases, there is no such link, and the appropriate contractual relationship involves a 'turnkey' contract with payment on completion of the construction phase.

2.2.3 Demand risk

Construction and operational risk relates primarily to the cost of the project. The most important factor affecting the benefits of the project is the demand for the services it generates. In most cases, demand in any given period may be represented by a downward-sloping demand curve, and the benefits generated by the project may be represented by the area under the curve. Projections of demand for the services of a given projection are subject to uncertainty. Moreover, this uncertainty tends to be greater the further into the future projections are made.

2.2.4 Network risk

The term 'network risk' describes a class of risks applying to an individual asset that is one part of a larger network, for example, an individual road in an urban road network. Usage of a particular road will depend, to a large extent, on decisions made with respect to other elements of the transport network. Hence, in many cases, it is inappropriate to consider the risks associated with an individual asset in isolation from the larger network.

2.3 The state-contingent framework

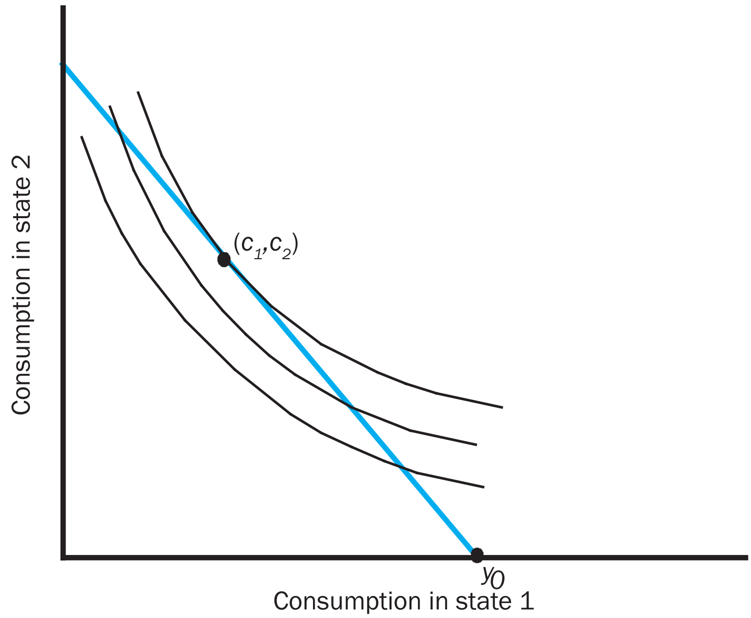
Uncertainty may be represented in a number of different ways. One common approach is to represent uncertain events by random variables with a given probability distribution. For many problems, this approach is mathematically convenient. However, in the analysis of decisions under uncertainty, such as the design and selection of projects, the random variable approach tends to obscure economically significant aspects of the problem. A slightly more elaborate approach, which brings out the similarities between choices under uncertainty and choices over time, is the state-contingent approach.

2.3.1 States of nature

The fundamental concept in the state-contingent approach is that of a 'state of nature'. A state of nature is a description of all the exogenous events relevant to the costs and benefits of a particular project. For example, in a road project, a full description of the state of nature might include weather conditions during the construction phase, the occurrence or absence of industrial disputes, growth in population in the areas served by the road and general economic conditions during the life of the road. Crucial features of the approach are that states of nature are exogenous and that the probability with which they occur is outside the control of those undertaking the project. Decisions made with respect to the management of a project will generate different outcomes in different states of nature. The outcomes, but not the states of nature themselves, can be managed and, to some extent, controlled.

The basic point is illustrated very simply in Figure 2. The only change from Figure 1 is that the axes now represent consumption levels in different states of nature, rather than in different time periods. In the presence of markets for state-contingent consumption, the line through y_0 may be regarded as a budget line, containing consumption vectors that may be traded for an amount y_0 received only if state 1 occurs.

FIGURE 2 CONSUMPTION CHOICES UNDER UNCERTAINTY



Using the state space approach, the outcomes of an uncertain decision may simply be represented as a list $y_1 \ldots y_S$ with one entry for each state of nature. As will be shown below, it is possible, in many cases, to attach a 'state-claim' price p_s to monetary payoffs arising in state s. Since the receipt of a dollar of income in every state of nature must be valued at a dollar, state-claim prices must satisfy

(13) $\Sigma\, p_s = 1.$

Given a set of state-claim prices, we may evaluate the uncertain outcome of a decision by the certainty equivalent

(14) $V = \Sigma\, p_s y_s.$

Except for the change in subscript, this is exactly the same as the present value formula. In Figure 2, the ratio of state-claim prices p_1/p_2 is given by the slope of the budget line, just as it would be if the axes represented consumption of different commodities or consumption at different points in time.

A particularly important case is that of risk neutrality, where the price p_s is simply the probability that state s will occur, denoted π_s, and we have

(15) $V = \Sigma\, \pi_s y_s$

$=E[\mathbf{y}],$

where $E[y]$ denotes the expectation or arithmetic mean of y. Obviously, the π_s satisfy (13).

The state-contingent approach encompasses the random variable approach. Given a list of possible states of nature 1 .... *S*, a random variable *x* may simply be specified by listing its outcomes x_1 .... x_S in each of the states. Note that two random variables might have the same probability distribution but yield different outcomes in different states of nature. That is, the random variable approach is a reduced form of the general state-contingent representation. The case of random variables with continuous distributions, such as the normal distribution, introduces a number of mathematical complications, but no fundamentally new ideas.

2.3.2 Event-tree analysis

A full description of a state of nature may be quite complex. In many cases, a compact description may be obtained using the notion of an event-tree, which describes a sequence of possible events over the life of the project. An event tree provides an integrated description of uncertain outcomes realised over time.

Example 2: Consider Example 1 presented in Section 1.2. We might consider three possible events that would affect the flow of income and consumption. First, we consider the possibility of a complete project failure. In event F_1 the project proceeds normally. In event F_2 the project is abandoned at the end of year 0 because of unexpected technical difficulties. Next, we consider two possible cost events. In event C_1 costs in year 1 are 400, as in Example 1. In event C_2 higher costs of 600 are incurred. Finally, we consider three possible demand events. In event D_1 demand grows at 2 per cent annually, as in the original example. In event D_2 demand grows at 4 per cent annually, and in event D_2 demand is stable.

There are 2x2x3 = 12 possible combinations of these events, which constitute the set of possible states of nature. In general, we might denote a typical state of nature as s_{121} consisting of the joint occurrence of events F_1, C_2 and D_1. In this particular case, however, the cost and demand events are relevant only if the project proceeds, so the set of relevant states of nature can be reduced to 7, as illustrated in Figure 3.

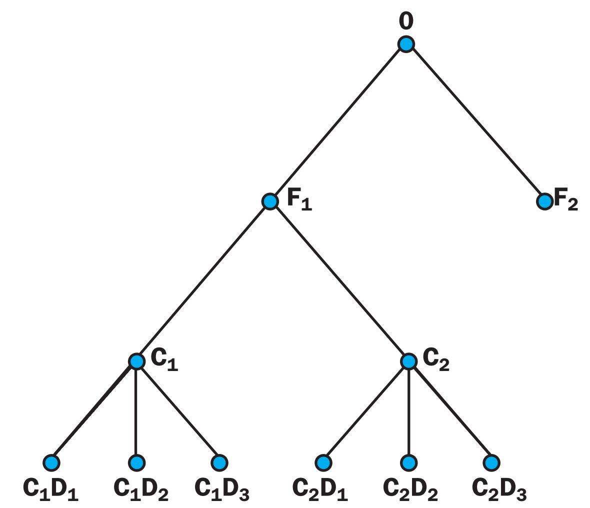
We will assume that all events are independent and that the probabilities are as follows:

(16) $\Pr\{F_1\} = 0.8$, $\Pr\{F_2\} = 0.2$;

(17) $\Pr\{C_1\} = 0.6$, $\Pr\{C_2\} = 0.4$;

(18) $\Pr\{D_1\} = 0.6$, $\Pr\{D_2\} = 0.2$, $\Pr\{D_3\} = 0.2$.

FIGURE 3 EVENT TREE FOR EXAMPLE 2



The information presented above is summarised in Table 2, where the annual net flows for each of the seven relevant states are presented in the columns 3–9. The first row contains the probabilities for the states. The data presented in Table 2 will be used as the basis for subsequent illustrative analysis.

Table 2 Net cash flows for example 2

Prob	*0.20*	*0.29*	*0.19*	*0.10*	*0.06*	*0.10*	*0.06*	
0	-300	-300	-300	-300	-300	-300	-300	-300
1	0	-300	-500	-300	-500	-300	-500	-304
2	0	100	100	100	100	100	100	80
3	0	102	102	104	104	100	100	82
4	0	104	104	108	108	100	100	83
5	0	106	106	112	112	100	100	85
6	0	108	108	117	117	100	100	87
7	0	110	110	122	122	100	100	88
8	0	113	113	127	127	100	100	90
9	0	115	115	132	132	100	100	92

2.4 Risk, probability and uncertainty

In the discussion thus far, the terms 'risk' and 'uncertainty' have been used in a fairly general sense. The terms 'risk' and 'uncertainty' are used in a variety of ways in general discussion of economic issues, and in some more technical ways by economists. It will be useful to define these terms more precisely.

A 'riskless' or 'certain' random variable x is one which has the same value $E[x]$ in every state of nature. Correspondingly the term 'pure risk' is used to describe variability of a state-contingent variable about the arithmetic mean. Among the many possible measures of riskiness are the variance and standard deviation.

2.4.1 Downside risk and variability

The technical usage of the term 'risk', described above, must be distinguished from a common interpretation of the term in ordinary use, to refer only to possible adverse outcomes. In the technical sense used here, risk arises both because the outcome may be less favourable than the mean, and because it may be more favourable.

When referring specifically to adverse outcomes, we will use the term 'downside risk'. Note that an adverse outcome must be defined relative to some notion of a 'normal' outcome, which need not be the arithmetic mean outcome.

2.4.2 Uncertainty as the absence of reliable probability distributions

Many of the more complex issues in the analysis of risk relate to individual judgements about unique events. Some decision theorists use the term 'risk' to apply to problems involving repeatable events with known objective probabilities, and uncertainty to apply to unique events where probabilities are subjective. Others restrict the term 'uncertainty' to cases where not even subjective probabilities are well-defined.

For the purposes of project evaluation, most of these controversies are not relevant. The analysis of projects under uncertainty requires the specification, implicit or otherwise, of probability distributions for the states of nature. These probabilities may be built up from observations on the relative frequency of the various events that characterise the state of nature. Alternatively, they may represent the informed judgement of those undertaking the analysis.

2.4.3 Representative values—mean, median and mode

In every elementary statistics course, students are presented with three concepts of the average. The first, and most commonly used is the (arithmetic) mean

(19) $E[x] = \Sigma \pi_s x_s$.

The second measure, also widely used, is the median or 50th percentile value. The median is useful as a measure of central tendency when the mean may be distorted by a small number of extreme values.

The third measure, typically mentioned once in introductory texts, and then forgotten, is the mode or most frequent value. Few useful statistical tools have been developed for dealing with modal values or variations about them. Nevertheless, modal values play a crucial role in intuitive thinking about averages and therefore about risk. For example, in considering the risks relating to some aspect of an infrastructure project, it is natural to focus on the modal, or most common, outcome and to consider risks as deviations from that outcome. In most cases, such deviations will be unfavourable. Hence, the mode will be higher than the mean or median value.

2.4.4 'No surprises' scenario and downside risk

The final outcome of an infrastructure project will depend on the realisations of a large number of random variables. One estimate of the outcome may be obtained by considering the result that would arise if all these variables took their modal values. This may be referred to as a 'no surprises' estimate. If for all or most variables the mode exceeds the mean and median, the 'no surprises' estimate obtained by cumulating a series of modal estimates will be greatly in excess of either the mean or the median.

Similar points apply to estimates derived using median values in cases where most variables are skewed to the left, so that the median is higher than the mean. The estimate of net present value for the entire project derived using median values for all variables will be higher than the mean or median net present value of the project.

Some of these points may be illustrated by considering Example 2. In this example, the median and mode coincide for each of the three events, and the 'no surprises' projection based on events F_1, C_1 and D_1 is the same as the original Example 1. However, Table 2 shows that the mean value of expected net benefits is less than that in Example 1 for every year.

If we combine the events F and C into a single event, with possible outcomes $\{F_2, F_1C_1, F_1C_2\}$, the modal event is F_1C_1 occurring with probability 0.48, but the median event is F_1C_2. More generally, combining median or modal projections may yield unpredictable results. By contrast, the mean has the appealing property of linearity. That is, for any random variables x, y, z and constants a, b, c

$$\text{(20)} \qquad E[ax+by+cz] = aE[x]+bE[y]+cE[z].$$

By virtue of this property, the mean value is unaffected by essentially arbitrary choices such as the decision to treat two events as distinct, or to combine them into a single event.

2.4.5 Variance and covariance

The concept of pure risk refers to variability of a random variable about the mean or some other measure of central tendency. A variety of measures of the variability or 'riskiness' of a random variable have been used or proposed. By far the most widely used is the variance

(21) $\sigma^2[x] = E[(x-E[x])^2]$.

Equivalently,

(22) $= E[x^2] - (E[x])^2$.

along with related measures including the standard deviation $\sigma[x]$, which is simply the square root of the variance, and the coefficient of variation

(23) $cv[x] = \frac{\sigma[x]}{E[x]}$.

The variance leads naturally to measures of the degree to which two random variables are related. For two random variables, x and y, the covariance is given by:

(24) $cov[x,y]= E[(x-E[x])(y-E[y])]$

Related measures include the correlation coefficient:

(25) $r(x,y) = \frac{cov[x,y]}{\sigma[x]\sigma[y]}$.

and the proportional covariance given by

(26) $\frac{cov[x,y]}{E[x]E[y]}$

Both the correlation coefficient and the proportional covariance are scale-independent. Although the correlation coefficient is valuable in statistical inference, the proportional covariance is more useful in the analysis of risk.

Another measure of association extensively used in the analysis of risk is the regression coefficient of y on x, given by

(27) $\frac{E[xy]}{E[x^2]}$.

This coefficient is commonly denoted by β in discussions of regression analysis.

The importance of covariance in project evaluation may be illustrated by considering the expected revenue of a project, which is given by

(28) $E[pz] = E[p]E[z] + cov[p,z]$.

Where prices are determined by a downward-sloping demand curve with supply a random variable, the covariance will be negative and expected revenue will be less than the product of expected price and expected output. More generally, it is important to note that, in most of the evaluation equations considered in benefit–cost analysis, covariance terms are negative.

A second crucial implication of covariance arises when risks are combined. Given two random variables x and y, the variance of their sum is given by:

(29) $\sigma^2[x+y]=\sigma^2[x]+\sigma^2[y]+2\mathrm{cov}[x,y]$.

Now suppose we combine an initial risk x with a small amount λ of an additional risk y. Then

(30) $\sigma^2[x+\lambda y]-\sigma^2[x]=\lambda^2\sigma^2[y]+2\lambda\mathrm{cov}[x,y]$.

If the amount of additional risk λ is small, the 'direct' impact $\lambda^2\sigma^2[y]$ can be disregarded, and the increase in variance will be determined primarily by the covariance $\mathrm{cov}[x,y]$ with the initial risk x. Note again that this term will enter evaluations as a negative impact.

2.4.6 Systematic and idiosyncratic risk

The concept of covariance may be used to explain one of the crucial ideas in finance theory, namely, that if analysis is conducted in terms of pure risk, only systematic risk (that correlated with aggregate income) is relevant. By contrast, idiosyncratic or diversifiable risk should not affect the evaluation of a project.

The basic point may be illustrated by supposing that all investors in a given market have the same risk attitudes and initially face the same risky distribution of income. (This is the equilibrium predicted by the standard efficient capital markets hypothesis if there are no differences in risk attitudes.)

Now consider two projects, A and B, with the same mean income and for both of which the variance of income is equal to σ^2. Suppose that the income from project A has positive covariance with existing income but that the income from project B has zero covariance.

Using equation (30), it can be seen that adding a small share λ of the income from project B to the portfolio of every investor makes only a negligible change $\lambda^2\sigma^2$ to the variance of total income. By contrast, adding a small share λ of the income from project A to the portfolio of every investor increases the variance of total income by an amount proportional to the covariance. Hence, other things being equal, a share in project B will be more valuable than a share in project A. The difference is due to the systematic risk associated with project B.

This discussion, in common with most standard financial analysis, assumes that investors can fully diversify all idiosyncratic risks and can achieve optimal risk-sharing for systematic risks. In fact, while investors can diversify their investment portfolios, they are, in general, faced by non-diversifiable risks which may be either systematic or idiosyncratic. For example, individuals with wage income face an uninsurable risk of income loss to unemployment which has both idiosyncratic components (such as the possibility of a dispute with an employer) and systematic components (such as the possibility of retrenchment in a recession). The implications of this point will be discussed in Section 4.

2.5 The expected utility model

Having formulated a description of a project involving uncertainty, it is necessary to consider how this description may be used in making a decision, in this case, whether or not to proceed with the project. A number of different models have been proposed to explain observed choices under uncertainty (positive models) and to suggest frameworks for optimal decision (normative models). The most prominent has been the expected utility model, which may be traced back to the work of von Neumann and Morgenstern (1944).

For any given consumption vector c, the expected utility model sets

(31) $V = \Sigma \pi_s u(c_s)$,

where π is a probability distribution, as before, and u is a utility function.

In the expected utility model, risk aversion is explained by concavity of the utility function u, which may be interpreted as diminishing marginal utility of consumption. That is, the larger is c_s, the smaller is $u'(c_s)$.

For small changes in consumption, the marginal utility vector is given by

(32) $\frac{\partial V}{\partial c_s} = \pi_s u'(c_s)$.

The analysis is unchanged by linear transformations of u. Since our object is to value projects in terms of units of present consumption, it is convenient to choose u so that the marginal utility of expected present consumption is equal to one

(33) $u'(E[\underline{c}])=1$.

If financial markets are sufficiently well-developed that a complete set of state-contingent consumption claims is freely tradeable, there will exist a well-defined state-claim price p_s for each s, as discussed in Section 2.3. If we normalise so that the price of a unit of certain income is 1, the state-claim price must be equal to $\pi_s u'(c_s)$ for all individuals. As in the standard general equilibrium framework under certainty, individuals trade until their marginal rates of substitution are equal.

The expected utility model involves preferences that are additively separable over states of the world. It is natural, though not strictly necessary, to combine this with preferences that are additively separable over time. Suppose now that each state of the world s corresponds to a stream of consumption flows $c_{s1} \ldots c_{sT}$. Then the objective function for the expected utility model with time-separable utility is:

(34) $V = \Sigma\Sigma \ \beta_t \pi_s u(c_{st})$

where β_t is a discount factor as before.

The associated state-claim prices, expressed in terms of non-stochastic present consumption, are given by

(35) $p_{st} = \beta_t \pi_s u'(c_{st})$.

That is, the price of units of consumption in state s, period t is the product of the discount factor for period t, the probability of state s and the marginal utility of consumption in that state (relative to that of nonstochastic consumption in period 0). Note that since $u'(c_s)$ is a decreasing function of c_s, the state-claim price for state s, period t will be lower, the higher is c_{st}. That is, because of the diminishing marginal utility of wealth, the covariance between state-claim prices and consumption levels is negative, just as in the revenue equation (28).

2.5.1 Constant relative risk aversion

In most project evaluation applications, the outcome should be independent of the scale of the project.

Suppose, for example, a given project involving improvements to a bridge yields positive net present value of, say, $1 million and that ten such projects, with identical and independent returns are available. Hence, each of the ten projects considered separately would yield net present value of $1 million. In the absence of funding constraints requiring the use of shadow prices, we would expect that combining the ten into a single large project would yield net present value of $10 million.

In the absence of risk aversion, this property of scale independence follows from the linearity of the expected value, which requires:

$E[k\mathbf{x}] = k\,E[\mathbf{x}]$ and,

(36) $E[\mathbf{x}+\mathbf{y}] = E[\mathbf{x}] + E[\mathbf{y}]$.

In expected utility theory, the property of scale-independence is referred to as constant relative risk aversion. The family of utility functions having this form is given by

(37) $u(c) = k\dfrac{c^{1-\gamma}}{(1-\gamma)} \qquad \gamma \neq 1$

$u(c) = k\ln(c) \qquad \gamma = 1.$

Note that

$u'(c) = kc^{-\gamma}$,

where γ is a parameter referred to as the coefficient of relative risk aversion and k is a scaling factor, which can be set equal to $E[c]^{-\gamma}$ to ensure $u'(E[c]) = 1$.

The coefficient of relative risk aversion plays a crucial role in determining the impact of risk on the evaluation of projects (or, more generally, of any risky project). To simplify the exposition, consider a single-period project, so that the issue of discounting may be ignored. Consumption in the absence of the project is denoted by $\underline{c}_0$. For a small project with risky return denoted by y, the value of the project may be expressed in terms of the certainty equivalent

(38) $\quad e = \Sigma\, p_s y_s$

which satisfies

(39) $\quad \mathrm{E}[u(\underline{c}_0 + e)] = \mathrm{E}[u(\underline{c}_0 + y)]$

Using a first order Taylor approximation around c and rearranging:

$$\mathrm{E}[y] + \frac{\mathrm{cov}\,[u'(c_0), y]}{\mathrm{E}[u'(c_0)]}$$

For the case of constant relative risk aversion we have a further approximation:

$$e = \mathrm{E}[y] - \gamma \frac{\mathrm{cov}[c_0, y]}{\mathrm{E}[c_0]}$$

Note that these approximations are valid only when y is small relative to c_0.

From the formula for the certainty equivalent, the risk premium is:

(40) $\quad \mathrm{E}[y] - e = \gamma \dfrac{\mathrm{cov}[c_0, y]}{\mathrm{E}[c_0]}$, or as a proportion: $\dfrac{\mathrm{E}[y] - e}{\mathrm{E}[y]} = \gamma \dfrac{\mathrm{cov}[c_0, y]}{\mathrm{E}[c_0]\mathrm{E}[y]}$.

Now the correlation coefficient between y and c_0 is, by definition, $r_{y,c_0} = \dfrac{\mathrm{cov}[c_0, y]}{\sigma[c_0]\,\sigma[y]}$. Substituting $\mathrm{cov}[c_0,y] = r_{y,c_0}\,\sigma[c_0]\,\sigma[y]$, and coefficients of variation $\mathrm{cv}[c_0] = \dfrac{\sigma[c_0]}{\mathrm{E}[c_0]}$ and $\mathrm{cv}[y] = \dfrac{\sigma[y]}{\mathrm{E}[y]}$, the risk premium as a proportion of $\mathrm{E}[y]$ is:

(41) $\quad \dfrac{E[y] - e}{E[y]} = \gamma \dfrac{\mathrm{cov}[c_0, y]}{\mathrm{E}[c_0]\mathrm{E}[y]} = \gamma\, r_{y,c_0}\, cv[c_0]\, cv[y].$

That is, the proportional risk premium is equal to the product of the coefficient of relative risk aversion and the proportional covariance. Given constant relative risk aversion, the proportional risk premium is independent of the scale at which the project is assessed.

2.5.2 Multiple goods

Thus far, the analysis has been based on the assumption that all flows of costs and benefits can be expressed in monetary terms, at given prices. Much of the literature on benefit–cost analysis is concerned with the problem of calculating appropriate prices for benefits such as reductions in travel time in order to permit evaluation of costs and benefits in terms of a common monetary unit.

The main concern in relation to uncertainty arises where a project supplies direct (that is, unpriced) consumption benefits to households (or intermediate inputs to business) and where a particular project supplies a large proportion of the total consumption of some particular good or service. An example is a water supply project that is the sole or primary source of water for a given community. If the quantity supplied is variable, a substantial risk premium may apply to the calculation of benefits.

The main issues may be illustrated with the aid of the simplifying assumption that the direct consumption benefits provided by a project may be summed up by a single service, denoted z, while consumption of all other goods is expressed in monetary terms by c, as before. Focusing on risk with respect to z, we may write an objective function

(42) $V = c + u(z)$

where u is a utility function. The implicit price of a marginal unit of z, expressed in monetary terms is simply $u'(z)$. If the expected level of services from a single-period project is given by $E[z]$, the benefits of the project are given by $E[u(z)]$. As in the case of monetary benefits, we have the approximation

(43) $E[u(z)] = u'(E[z])E[z]\left(1 - \frac{\gamma\sigma^2[z]}{E^2[z]}\right).$

This differs from (41) by virtue of the inclusion of the expected price $u'(E[z])$, and the use of the variance of z rather than the covariance between y and z.

2.5.3 Critiques of the expected utility model

Until the early 1980s, the expected utility model dominated economic analysis of choice under uncertainty and over time. Since then, a wide variety of alternatives and generalisations have been proposed. The motivation for most of these alternative approaches has been an accumulation of evidence that individuals faced with choices involving time and uncertainty do not generally make choices consistent with the predictions of expected utility theory. Allais (1953) was the first to design a choice problem in which individuals generally did not conform to the predictions of expected utility theory. Similar examples involving choice over time have been put forward by Loewenstein and Thaler (1989).

The first set of alternative models, including those of Kahneman and Tversky (1979), Chew (1983), Machina (1982) and Quiggin (1982) focused on choice under uncertainty, and relaxed the assumption of an objective function that is linear in probabilities. More recent work has focused on choice over time, and has been concerned with models of hyperbolic discounting, in which discount factors decline more rapidly for events in the near future than for those in the distant future.

These generalised models have accounted for a range of important economic phenomena, such as participation in lotteries. Nevertheless, the expected utility framework remains central to benefit–cost analysis. There are two reasons for the continuing reliance on expected utility theory. First, despite its limitations as a descriptive model, the expected utility approach has strong normative appeal, and benefit–cost analysis is a normative task.

Second, the additivity property of the expected utility framework is essential if project appraisal is to be undertaken in an analytical fashion, with individual components being evaluated and these evaluations aggregated to produce a comprehensive evaluation. An example of this analytic approach is the standard procedure of 'rolling back' decision trees in the analysis of complex choices under uncertainty (Lavalle and Wapman 1986). For these reasons, analysis in this paper will be confined to the expected utility framework.

2.5.4 A simple example

The state-contingent approach to the evaluation of risky projects may be illustrated using a simple example. We will separate the issues of discounting and risk by assuming that all benefits from a project are received in a single time period. The cash flows from period 9 in Example 2, reproduced below, will be used for illustration.

Prob	0.20	0.29	0.19	0.10	0.06	0.10	0.06
Benefits	0	115	115	132	132	100	100

We first consider the case when the flow of benefits is expressed in monetary terms, say, millions of dollars. It must therefore be evaluated in terms of its covariance with income in the absence of the project. It seems reasonable to assume that income in the absence of the project is uncorrelated with cost conditions for the project, but may covary with demand conditions. We will assume that, in the absence of the project , income takes the value \$100 billion in demand event *D*1, \$110 billion in demand event *D*2 and \$90 billion in demand event *D*3.

Since the ratio of project benefits to total income is around 0.1 per cent, the project is 'small' relative to total income, and its direct contribution to the variance of total income may be disregarded. As this example, indicates even projects that are quite large in absolute terms will be small in this sense. Attention may therefore be focused on the covariance term. The proportional covariance may be computed as approximately 0.5 per cent. Hence, for the case of constant relative risk aversion, equation (41) shows that the proportional risk premium is 0.005γ, where γ is the coefficient of relative risk-aversion. For plausible values of γ, in the range 0 to 2, the certainty equivalent return for the project is computed by deducting a risk premium of between 0 and 1 per cent from the mean return. Since the mean return is around \$92 million, the certainty equivalent return is between \$91 million and \$92 million.

The risk premium in this case is small. Consider, however, the case when the project is a major source of a particular service. Suppose, for example, that the project is a water supply project, output is measured in megalitres (ML) per user, the output price is \$10/ML and that, in the absence of the project, users will receive 50 ML each. Suppose that the expected supply from the project is 500 ML per user, but that there is a 10 per cent chance that the project will fail and deliver no water. The mean output is 450 ML per user, the standard deviation is 158 ML and the coefficient of variation is 0.31. The approximate estimate of the risk premium from equations (23), (25) and (40) is $450*(0.31)^2\gamma = 550\gamma$ where γ is the coefficient of relative risk aversion. Thus, for $\gamma=2$, the risk premium is greater than \$100 per user, implying that a project with a riskless output of 350ML would be preferred. (An exact calculation yields a premium that is slightly larger, around 135 in this case). Thus, for projects that deliver a large proportion of the total supply of some particular service, such as water supply, or transport between two points, the risk premium may be substantial.

3 Risk and project evaluation under uncertainty

3.1 Introduction

There is a close analogy between project evaluation under uncertainty and discounting under certainty. This does not mean, however, that uncertainty can be dealt with by adjusting the rate of discount that would be used under certainty. As will be shown below, this procedure is appropriate only in special cases.

As was shown in Section 1, the basic objective in discounting is to convert flows of payments and returns over time into a single present value, by expressing increments to consumption at different points in time in terms of a common 'currency'. Similarly, in project evaluation under uncertainty, the central objective is to aggregate flows in different states of nature by expressing them in terms of a common currency, commonly called the certainty equivalent.

Uncertainty is typically resolved over time. The pricing approach set out above permits the evaluation of uncertain income flows received over time. The starting point is a set of contingent income flows $y(s,t)$ where $y(s,t)$ is the income received in period t under state of nature s. Under the additivity assumptions that are standard in the expected-utility model these income flows may be aggregated in two stages: over states of nature and over time. The final answer is independent of the order of aggregation.

The analysis in Section 1 shows how a flow of income received over time may be converted into a single present value, representing the equivalent amount of income received at time 0. The analysis in Section 2 shows how an uncertain income, received at a given time *t*, may be converted into a certainty equivalent of equal value.

A natural two-stage procedure is to deal with uncertainty first, and then evaluate flows over time. In the first stage of this procedure, the state-contingent income flows realised in period *t* are aggregated to determine a certainty-equivalent income flow. This is a sequence of non-stochastic flows, y_0 ... y_T which yield the same expected utility in each period as the state-contingent flows under consideration. In the second stage, the certainty-equivalent flows are reduced to a present value by the usual discounting procedure. Since these flows are non-stochastic, it is clearly appropriate to use the riskless discount rate.

It is equally possible, though perhaps less intuitive in the context of present value, to deal with time first, and then take account of uncertainty. That is, for each possible state of nature, the present value of the stream of income conditional on that state of nature occurring may be computed. The expectation taken over all states of nature yields the certainty-equivalent present value.

The basic ideas may be illustrated using Example 2. We will first consider the case of risk-neutrality, so that state-contingent prices are given by probabilities. Now the two approaches are represented in Table 3, which consists of the data for Table 2 with the addition of an initial row and column showing the probabilities and discount rates respectively and a final row and column showing net present values and means respectively. The expected net present value may be obtained either by calculating the NPV for the final column or calculating the mean of the final row. The result, given in the bottom right-hand corner is negative.

Thus, whereas the 'no surprises' projection *C*1*D*1 yields a positive NPV, the mean value obtained when uncertainty is taken into account is negative.

3.2 Implementing the state-contingent approach

As the discussion of Example 2 has shown, the state-contingent approach to project evaluation involves three fundamental requirements: an appropriate specification of the state space; the derivation of accurate estimates of state-contingent distribution of income; and the use of appropriate prices for state-contingent income flows.

3.2.1 Specifying the state space

A number of steps may be taken towards this goal. First, it is desirable to allow for possible technological shocks, defined broadly to cover events affecting the costs of a project. These may be particular adverse events that arise with positive

Table 3 **Project evaluation under uncertainty: risk-neutral case**

0	1.00	-300	-300	-300	-300	-300	-300	-300	-300
1	0.95	0	-300	-500	-300	-500	-300	-500	-302
2	0.91	0	100	100	100	100	100	100	80
3	0.86	0	102	102	104	104	100	100	82
4	0.82	0	104	104	108	108	100	100	83
5	0.78	0	106	106	112	112	100	100	85
6	0.75	0	108	108	117	117	100	100	87
7	0.71	0	110	110	122	122	100	100	88
8	0.68	0	113	113	127	127	100	100	90
9	0.64	0	115	115	132	132	100	100	92
NPV		-300	71	-119	116	-74	30	-161	-61

probability in projects of a given class, such as the discovery of adverse geological conditions in tunnel projects. Alternatively, the shocks may be described more generically, for example, as a class of events affecting the date on which a project is completed.

Second, it may be useful to specify distributions for input and output prices. Assuming these are exogenous to the project, the prices themselves may be treated as defining the states of nature in which they occur. Similarly, it is desirable to specify events relevant to the demand for the services of the project.

3.2.2 Estimating income flows

The next step in the process is the specification of income flows for each possible state of nature, represented by a value for each of the relevant events. It is natural to do this using the event-tree structure illustrated in Figure 3. Note that if an event is not realised until time t, the income streams for states of nature that differ only with respect to that event must coincide at least until time t-1. In Example 2, the rate of growth of demand has no effect until year 3, so that income streams for states *C1D*1, *C1D*2 and *C1D*3 must coincide for years 1 and 2, and similarly for *C2D*1, *C2D*2 and *C2D*3.

For simple cases such as that of Example 2, the income flow associated with each state of nature may be computed manually or using a spreadsheet. In more complex cases, it may be appropriate to use simulation models or special-purpose tools such as the computer program @Risk.

3.2.3 State-contingent prices

Having defined the states of nature and estimated the state contingent flows, it is necessary to determine the state-contingent prices to be used to value flows of the form $y(s,t)$ in terms of non-stochastic income. In the expected utility model, two parameters are relevant. The first is the probability with which a given state occurs. The second is the marginal utility of income in that state, which will depend, in general, on the level of consumption.

For risk-averse decision-makers, the value of state-contingent income depends on covariance with existing income. For an individual with an initial income stream c, a small additional flow y is valued at

(44) $$\begin{aligned} V &= \Sigma_t\Sigma_s\, p_{st}y_{st} \\ &= \Sigma_t\, \beta_t\, E[\, u'(c_{st})y_{st}] \\ &= \Sigma_t\, \beta_t\, \{E\,[u'(c_{st})]E[y_{st}]+\mathrm{cov}[u'(c_{st}),y_{st}]\}. \end{aligned}$$

(Note that taking expectations is the same as taking a probability-weighted sum over the states of nature.)

Assuming a constant real interest rate r, at which the individual can trade freely, equilibrium requires

(45) $$\begin{aligned} p_t &= \beta_t\, \{E\,[u'(c_{st})]\} \\ &= (1+r)^{-t}. \end{aligned}$$

and substitution into (44) yields

(46) $$V = \Sigma_t\, \{p_t\, E[y_{st}] + \beta_t \mathrm{cov}[u'(c_{st}),y_{st}]\}.$$

The impact of systematic risk arises from the covariance between the state-contingent price of income, which is given by the marginal utility of state-contingent income $u'(c_{st})$ and the flow of returns from the project y_{st}. This covariance will normally be negative. Since pure idiosyncratic risk has, by definition, zero covariance with consumption, it is disregarded in project evaluation.

Equation (46) provides the basic rationale, in terms of expected utility, for applying a lower valuation to projects characterised by systematic risk. The risk premium arises from the covariance between the marginal utility of consumption and the flow of services provided by the project.

From (45), the first term in the valuation (46) corresponds to the net present value of mean income. Notice that, even if there is no inherent discounting, so that $\beta_t = 1$, the equilibrium interest rate *r* will be positive as long as consumption tends to rise over time, so that $E[u'(c_{st})] < u'(c_0)$.

For the case of preferences with constant risk aversion we can simplify further, obtaining

(47) $\quad \mathrm{cov}[u'(c_{st}), y_{st}] = \mathrm{cov}\,[c_{st}^{-\gamma}, y_{st}]$

which is negative if project returns y_{st} are positively correlated with consumption growth c_{st}.

3.2.4 Projects providing services directly to consumers

The discussion above has focused on the case where all costs and benefits from the project are expressed in monetary terms. Using the approach set out above, it is straightforward to deal with the case where the project yields unpriced services to consumers. All that is required is a reinterpretation of the terms used in the previous section. The flow of services provided by the project, *y*, should now be interpreted as units of the consumption good, and the initial consumption level *c* refers to consumption of this good, rather than to total consumption.

The interpretation of the risk premium as the covariance between the marginal utility of consumption and the flow of services provided by the project remains unchanged. However, in the case of a single good, such as water supply, it need not be the case that the flow of services provided by the project is small in relation to total consumption. Hence, as in the example of section 2.5.4, it is necessary to take account of both variance and covariance terms. The variance encompasses both systematic and idiosyncratic risk, while the covariance is derived solely from systematic risk.

A related case is that where the project provides intermediate inputs to production processes. A road project, for example, will provide both direct benefits to households and reductions in transport costs for firms. These are normally valued by estimating savings in vehicle operating costs and time. Work time is valued at the wage rate and non-work time at a proportion of the wage rate derived from studies of data from situations where people trade off time against costs. The amount of traffic benefiting from the road project, and hence the total benefits from the project, may vary with aggregate demand, implying that it is necessary to take account of systematic risk.

In addition, where the costs of the project are met from taxation revenue, while the benefits accrue to households or firms, it is necessary to take appropriate account of the deadweight costs of raising revenue, as discussed in Section 1.

3.3 *Further analysis of Example 2*

Adjustments for 'pure' risk may be illustrated using Example 2. Preferences are assumed to be of the form $u(c) = \ln(c)$, that is, constant risk aversion with $\gamma=1$. For simplicity, it will be assumed that there is no inherent time preference, that is, $\beta_t=1$ for all t. Discounting of marginal future consumption therefore arises exclusively from the fact that, on average, consumption levels are increasing over time.

To maintain comparability with Table 3, growth rates for consumption will be selected so that the discount factor $E[u'(c_{st})]$ is approximately 0.95, consistent with a real interest rate of 5 per cent. It will be assumed that the central demand event *D*1 arises when aggregate consumption grows at a rate of 5 per cent per year, that the high demand event *D*2 arises when aggregate consumption grows at a rate of 10 per cent per year and that the low demand event *D*3 arises when aggregate consumption is constant. Thus, in any time period, the flow of services generated by the project is positively correlated with aggregate consumption. This implies that the value of the project must be adjusted downwards to take account of risk aversion.

The state-contingent prices, $\beta_t u'(c_{st})$ for events *D*1, *D*2 and *D*3 are shown in Table 4, along with the discount factor applicable to a non-stochastic payment, which is given by $E[u'(c_{st})]$.

Table 4 State-contingent prices for example 2

0	1.00	1.00	1.00	1.00
1	0.91	0.86	0.95	0.91
2	0.82	0.73	0.91	0.82
3	0.75	0.63	0.86	0.75
4	0.68	0.54	0.82	0.68
5	0.61	0.46	0.78	0.62
6	0.56	0.40	0.75	0.56
7	0.51	0.34	0.71	0.51
8	0.46	0.29	0.68	0.47
9	0.42	0.25	0.64	0.43

The meaning of Table 4 may be illustrated by an example. A unit of income received in period 4 under demand scenario 2 has a risk-adjusted present value of 0.54 units. This value reflects the discounting applied to nonstochastic income, under which a unit of nonstochastic income in period 4 has a present value of 0.68 units and the fact that demand scenario 2 is one of higher-than-average consumption, and therefore lower-than-average marginal utility of income.

A risk-adjusted version of the analysis presented in Table 3 may now be considered. The results are shown in Table 5. Columns 2 to 8 show the flows of consumption considered previously, multiplied by the appropriate state-contingent prices from Table 4. Column 9 shows the mean return discounted using the non-stochastic discount factor from the final column of Table 4. Column 10 shows the risk adjustment required to take account of the covariance between aggregate consumption and returns from the project. Column 11 shows the net return.

The most striking feature of Table 5 is that the risk adjustment is quite small—equal to about 1 per cent of the present value of total project benefits. The adjustment would be greater if a higher risk aversion parameter were used, but most studies suggest that $0 \leq \gamma \leq 2$, so that even a high value of γ would only raise the adjustment from 5 to 10. Similarly, the premium would be larger if aggregate consumption were more risky but the variability used in the example is considerably greater than that observed in practice. Finally, the adjustment would be larger if the project displayed more systematic risk in returns. Once again, however, even if the project were twice as risky, the risk adjustment would still be modest.

For this example, the adjustment for pure risk is likely to be between 5 and 20. Recall, by contrast, that whereas the 'no surprises' analysis in Table 1 yielded a positive NPV, the state-contingent analysis of Table 2, taking appropriate account of downside risk, reduced net benefits by 134, from +71 to -63.

Table 5 Project evaluation under uncertainty: Risk averse case

Year	*F2*	*C1D1*	*C2D1*	*C1D2*	*C2D2*	*C1D3*	*C2D3*	*Mean*	*Covar*	*Net*
0	–300	–300	–300	–300	–300	–300	–300	–300.0	0.0	–300.0
1	0	–285	–475	–270	–450	–300	–500	–289.5	0.0	–289.5
2	0	90	90	81	81	100	100	72.6	0.0	72.6
3	0	87	87	76	76	100	100	70.5	–0.1	70.4
4	0	85	85	71	71	100	100	68.5	–0.2	68.2
5	0	82	82	66	66	100	100	66.5	–0.4	66.1
6	0	80	80	62	62	100	100	64.7	–0.6	64.1
7	0	77	77	58	58	100	100	62.9	–0.9	62.0
8	0	75	75	54	54	100	100	61.1	–1.2	59.9
9	0	72	72	51	51	100		59.4	–1.5	57.9
Total								–63.3	–5.0	–68.3

3.4 Practical issues

A number of practical issues arise in implementing the state-contingent approach. First, there is the choice between representing distributions in discrete terms, as has been done here, or in terms of continuous distributions, such as the normal distribution and various related distributions (Student's *t*, *F* distribution and so on). The main advantage of the latter approach is that there exists a well-developed statistical theory for normally distributed variables. Using this theory, it is possible to derive standard errors and confidence intervals for estimates. In addition, many commonly-occurring stochastic processes give rise to variables with normal distributions. In project evaluation, however, these advantages are not, in general, sufficient to offset the increased complexity and reduced transparency of procedures involving continuous distributions.

Assuming that discrete distributions are used, it is necessary to consider how disaggregated they should be. In the case of demand growth, for example, it is clearly possible to specify probability distributions arbitrarily finely. However, in many cases there is little benefit from considering more than three possible outcomes. A three-point distribution allows for the mean, variance and skewness of the distribution to be specified independently. Only in cases where the outcome of the project depends critically, or in a highly nonlinear fashion, on a particular uncertain variable is a more detailed specification likely to yield substantial additional precision.

3.4.1 Scenarios and sensitivity analysis

Explicit representation of uncertainty has not been usual in project evaluation. To the extent that uncertainty has been taken into account at all, the most common approach has been the presentation of a number of analyses incorporating different assumptions. This procedure is commonly referred to as 'sensitivity analysis', particularly when the analyses differ from the base case by virtue of variations in one or more parameters, such as the rate of growth of demand in Example 2. In cases where the alternative assumptions do not have a natural parametric representation, they are sometimes referred to as 'scenarios'. The possible events affecting costs in Example 2 might be described in this way.

Sensitivity analyses are not, in general, accompanied by probabilities, nor are they supposed to give a complete representation of the possible events relevant to the project evaluation. Given the central role of uncertainty, the simplification obtained by not presenting explicit probabilities is, in most cases, more than offset by the potential error that is introduced through a focus on a single projection that may be based on median or modal values.

The best use for sensitivity analysis is to consider the impact of alternative discount rates on the project evaluation. These variables are external to the evaluation itself, and those undertaking the project evaluation are unlikely to have special information about likely future discount rates or about the shadow prices applicable in situations where the total volume of investment is constrained.

3.5 Risk-adjusted discount rates as an approximation

In practice, with the occasional exception of sensitivity analysis, very few applications of benefit–cost analysis incorporate a state-contingent analysis of the type suggested in the previous section. The most commonly-used procedures are based on evaluation of a non-stochastic projection in which uncertain variables are estimated using some form of representative value. Risk is either ignored or 'taken into account' by using a higher discount rate.

As has already been noted the problem of evaluating uncertain outcomes over time is one of expressing, in terms of present income, the price of income flows received in some future period, and in a specific state of the world. In an analysis based on a single mean value estimate for each future time period, the average state-contingent price associated with that period is, in effect, a weighted average of the prices attributable to income flows for all possible states of the world in that period. The riskier the income in a given period, and the higher the systematic component of that risk, the lower should be the average state-contingent price for that period.

The use of a higher discount rate to evaluate risky investments involves attributing a lower average state-contingent price to income flows at later dates. Intuitively speaking, this approach will approximate the correct answer if the riskiness of income flows is greater the further in the future they are received.

This section will present a critical assessment of the practice of applying a risk-adjusted discount rate to a 'central projection' of project outcomes. Circumstances under which this procedure yields the same outcome as the ideal approach will be described, as will circumstances when the outcomes are likely to be seriously misleading.

3.5.1 CAPM and the appropriate use of risk-adjusted discount rates

Cases where it may be appropriate to use a risk-adjusted discount rate arise in the context of finance theory, and particularly in intertemporal versions of the Capital Asset Pricing Model (CAPM).

In a single period-setting, CAPM begins with the idea that if capital assets can be freely traded in a securities markets then any idiosyncratic risk can be eliminated by diversification. Hence, only systematic risk is relevant in the evaluation of capital assets such as corporate securities (stocks, bonds and so on). Supposing that the initial portfolio x is a representative sample of the market as a whole, the argument leading to equation (27) shows that the change in the variance associated with a small holding λ of some asset y is determined by the covariance

(48) $\sigma^2[x+\lambda y] - \sigma^2[x] = \lambda^2\sigma^2[y]+2\lambda \mathrm{cov}[x,y]$

$\approx 2\lambda \mathrm{cov}[x,y]$.

In finance theory, it is more usual to express the relationship in terms of rates of return using the coefficient arising from a regression of returns to a particular investment on returns to the market as a whole. This coefficient is often called the 'beta' coefficient, but we will denote it by *b* to avoid confusion with the discount factor defined above.

For the single period case, we obtain

(49) $r = r_0 + \beta r_m$

where r is the rate of return required for the project, r_0 is the risk-free rate of interest, β is the regression coefficient and r_m is the rate of return for the market as a whole. The term βr_m may be interpreted as a risk adjustment for the discount rate.

This pricing relationship may be extended to a multiperiod equation under the assumption that returns to capital assets follow a random walk. This implies that the proportional risk associated with holding such an asset grows linearly over time.

The use of a risk-adjusted discount rate for evaluating investment projects is defensible if and only if conditions analogous to those for the multiperiod version of CAPM are satisfied, or at least approximately satisfied. The main conditions are:

(i) all projects should have similar risk characteristics (that is, similar values of β);

(ii) there should be a single investment period, followed by a period of positive income flows (the same for all projects) with constant mean income flow; and

(iii) the variance of income flows should increase linearly over time.

These conditions can be relaxed and generalised in various ways. For example, condition (iii) can be generalised to allow for some risk index other than the variance to increase linearly over time provided preferences are appropriately linear in this index.

3.5.2 Inappropriate use of risk-adjusted discount rates

Having considered conditions under which the use of a risk-adjusted discount rate may be appropriate, it is easy to see why this procedure will frequently be inappropriate. In most project evaluation problems, one or more of the necessary conditions is not satisfied. These will be considered in turn.

It is obviously inappropriate to make an identical risk adjustment to discount rates if projects with different risk characteristics are being compared. There are, however, some procedures that can deal with this problem. If riskiness can be summed up using an intertemporal version of the capital asset pricing model, then each project can be attributed a discount rate based on the beta coefficient associated with its income flows. Note that the intertemporal CAPM incorporates the assumption that income flows follow a diffusion process with parameters such that risk increases linearly over time.

The problem of timing is less tractable. As with the use of internal rates of return, the procedure breaks down completely if costs, rather than benefits, are incurred in later stages of the project. The use of a higher discount rate in cases of this kind has the effect of making risky projects appear more, rather than less, attractive, since the (negative) present value of future costs is made smaller. More generally, because the procedure of adjusting discount rates confounds the price of risk and the price of time, it works only if all projects have similar timing characteristics.

Finally, even where the first two conditions are satisfied, risk may not grow linearly over time. To the extent that uncertainty primarily relates to estimates of demand for the services provided by a project, it seems reasonable that such uncertainty should be greater, the further in the future projections are made. This point is illustrated by the analysis of Example 2, presented in Section 3. In this case, the only systematic risk arises with respect to demand, and the risk grows linearly over time. As a result, the risk adjustment can be approximated quite closely by raising the discount rate from 5 per cent to 5.2 per cent. However, where the relationship between risk and aggregate consumption is more complex, or where different projects are being compared, this simple adjustment will not work.

3.6 Dealing with biased estimates of project costs and benefits

A second possible justification for the use of a risk-adjusted discount rate is that the discount rate may compensate for biases in estimates of average returns. As has already been noted, such biases are likely to arise if modal or 'surprise-free' representative values are used in the analysis. In such cases, estimates are likely to involve considerably more 'downside' than 'upside' risk.

There is a good deal of evidence to suggest that *ex ante* project evaluations yield estimates that are biased in favour of acceptance of the project. Pohl and Mihaljek (1992) show that the mean estimated *ex post* return on a sample of World Bank projects was 16 per cent compared to an average predicted return of 22 per cent (medians were 14 and 18 per cent). Even the estimated *ex post* returns are not definitive, and may be over-estimates in some cases.

Results of this kind are evidence of the need for careful consideration of possible adverse events. In addition, more systematic *ex post* assessment of projects would clearly be helpful. However, the most common response has been to adjust discount rates.

Although the use of an adjusted discount rate may be defended as an approximation in dealing with, for example, pure demand risk, the use of discount rate adjustments as a way of taking account of 'downside risk' is a dangerous and erroneous practice. A number of problems arise when this practice is adopted.

The first problem is that, unlike the case of demand risk, there is little reason to suppose that downside risk will grow linearly over the life of a project. One major source of downside risk is the possibility of unforeseen difficulties in the construction phase of the project. A second is that, on completion, the services delivered by the project will not be of the quality and volume expected. Alternatively, it may become apparent that the service specifications were inappropriate.

Consider as an example a toll road project involving tunnels and bridges. Construction costs may be increased and completion of the project delayed by unfavourable geological conditions leading to leaks in tunnels or unstable foundations for bridges. If *ex ante* project evaluations do not take account of this possibility, they will be subject to downside risk. Upon completion, the tolling system may work imperfectly, or generate consumer dissatisfaction. Alternatively, the project may be technically successful, but drivers may seek to avoid tolling points, thereby creating congestion on connecting roads. Again, if these possibilities are not taken into account, downside risk will arise.

As can be seen from these examples, many of the sources of downside risk are likely to arise or become apparent early in the life of a project. Projects with high risks of this kind are not penalised by the adoption of a high discount rate.

By contrast, projections of demand growth over time are typically derived from statistical procedures that are subject to pure risk, but are designed to minimise bias when applied appropriately. Hence, there is no reason to suppose that they will be a source of additional downside risk.

It follows that adjustments to discount rates will not, in general, provide an appropriate way of offsetting downside risk. In comparing projects with similar risk characteristics, the procedure of adjusting discount rates will inappropriately favour projects with short-lived benefits as well as those with long-lived costs.

A second problem relates to incentives in the project evaluation procedure. If discount rates are adjusted to compensate for the failure to take appropriate account of downside risk, projects for which such risks are taken into account will receive less favourable evaluation. Hence, if those undertaking project evaluation wish to ensure a socially optimal project selection, they must adopt procedures that leave some risks unaccounted for. There is only a short step from this situation to one where the project evaluation process becomes one of advocacy, with project proponents putting forward unrealistic projections as the basis of analysis, and central agencies counteracting this bias by raising discount rates.

Problems arising from errors in ex ante estimates of project costs and benefits can never be avoided completely. However, the approach suggested in Section 3 is more appropriate than that of adjusting discount rates. This procedure will eliminate the bias against projects with long-lived benefits and the bias in favour of projects with long-lived costs.

4 The cost of pure risk

As has been shown above, the value of a risky project may be regarded as the difference between the expected present value of the project and the 'pure' risk premium arising from covariance between project returns and aggregate consumption.

In this section, it will be argued that, in many cases, the effect of the risk premium is zero or negligible. Hence the appropriate procedure is simply to calculate expected NPV using the real bond rate.

4.1 CAPM and CCAPM

In most applications of the CAPM model, the price of risk is taken as a 'free' parameter, that can be estimated to fit market data. However, as is shown by the discussion in Sections 2 and 3, the economic analysis underlying the CAPM model implies that the price of risk must be determined by the covariance between the return to the asset and the marginal utility of income, normalised in terms of period 0 income.

The full CAPM model in which the price of risk is determined endogenously, is referred to as the Consumption-based Capital Asset Pricing Model or CCAPM. The free parameter in equation (49) is r_m, the rate of return on the market portfolio. The rate of return on the market portfolio may also be expressed in terms of the (relative) market risk premium r_m-r_0, which is related to the price of risk by the general relationship

(50) $$r_m - r_0 = \frac{-\mathrm{cov}\,[u'(c_s),Y_s]}{\mathrm{E}[Y_s]}\ .$$

where Y_s is the total return to the market portfolio in state s.

Estimation of the price of risk is simplified by the assumption of constant relative risk aversion. Using equations (37) and (40), the expected utility model with constant relative risk aversion implies that the relative risk premium should be approximated by

$$(51) \qquad \frac{(E[y]-e)}{E[y]} = \gamma \frac{\text{cov}\,[c_s, Y_s]}{E[c_s]E[Y_s]}$$

$$(52) \qquad = \gamma \frac{r(c_s, Y_s)\,\sigma(Y_s)\,\sigma(c_s)}{E[c_s]E[Y_s]}$$

The relative standard deviation of aggregate consumption averaged over time is around 3 per cent. That is $\sigma(c_S)/E[c_S]= 0.03$. It might reasonably be assumed, then, that the systematic component of the coefficient of variation of the benefits of a typical project (that is the component correlated with consumption) is also around 3 per cent, so that $r(c_S, Y_S)\ \sigma(Y_S)\ /E[Y_S] = 0.03$. Substituting these values into (52) yields

$$\frac{(E[y]-e)}{E[y]} = \gamma 0.03 \times 0.03 = 0.0009\gamma$$

Hence, if the coefficient of relative risk aversion is equal to 1, the risk premium, expressed as a proportion of the expected flow E[*y*] should be approximately 0.1 per cent. That is, a project with a benefit–cost ratio of 1.001 or more, evaluated on the assumption of risk-neutrality, would still have a benefit–cost ratio greater than one (that is, positive NPV) after taking account of risk. Given the imprecision with which critical parameters such as the real bond rate are known, the CCAPM model implies that there is little benefit in taking account of risk.

4.2 The equity premium puzzle

Understanding of these issues may be enhanced by analysis of the ‘equity premium puzzle’. Long data series generally show that the rate of return to buying and holding the market portfolio of stocks is considerably greater than the rate of return to government bonds. For example, Mehra and Prescott (1985) present data showing that over the period 1889–1978, the average annual yield on the Standard and Poor 500 Index was seven per cent, while the average yield on short-term debt was less than one per cent. Using the model of intertemporal optimisation of consumption derived above, and evidence on the growth and variability of aggregate consumption, Mehra and Prescott compute equilibrium asset prices for debt and equity under a wide range of parameter values. They show that the equity premium should be no more than half a per cent.

Mehra and Prescott coined the term ‘equity premium puzzle’. The observed data constitutes a ‘puzzle’ because it seems to suggest that individual investors are not rationally optimising and also that there are unexploited opportunities for arbitrage. Risk aversion does not seem an adequate explanation; although individual shares are risky, diversification should reduce risk greatly.

4.3 Explanations of the equity premium puzzle

Many explanations of the equity premium puzzle have been offered. For the purposes of the present paper, it is useful to divide these into three broad categories. First, there are explanations consistent with the efficient markets hypothesis, and some version of CCAPM. Second, there are explanations that depend on specific real or perceived characteristics of equity as an asset. Third, there are explanations that rely on transactions costs or other forms of market failure.

The simplest explanation of the equity premium, consistent with rational optimisation and efficient markets, is that the representative consumer is many times more risk-averse than is normally supposed. This explanation has not found much support in view of extensive evidence supporting the view that the typical coefficient of relative risk aversion is close to 1.

The most plausible explanation of the equity premium in terms of the characteristics of equity as an asset is that of Swan (2000), based on the greater liquidity of bonds. If correct, this explanation would imply that the correct discount rate for public projects is the rate of interest on bonds.

Other proposed resolutions of the equity premium puzzle include models of preferences with habit persistence (Constantinides 1990) or dependence on previous peak consumption. However, no satisfactory explanation of the equity premium puzzle has been found that is consistent with the efficient markets hypothesis.

The most important evidence against the idea that the observed equity premium arises from efficient markets is based on comparisons between the variability of individual consumption and the variability of aggregate consumption. As Heaton and Lucas (1996) and Constantine, Donaldson and Mehra (1998) observe, the relative standard deviation of individual consumption is around 20 per cent, far greater than the 3 per cent variation in aggregate consumption. This implies that the extent to which risk in individual consumption is diversified is considerably less than would be the case if the efficient markets hypothesis were valid. As Weil (1989) notes, under standard assumptions about preferences, individuals will be less willing to invest in risky equities in the presence of non-systematic 'background risk'.

Among explanations that focus on real or perceived characteristics of equity, the most notable are the claims that investors overestimate the riskiness of equity (Glassman and Hassett 1999) or that investors prefer bonds to equity because they yield liquidity benefits (Swan 2000). Both explanations are consistent with the empirical claim that the equity premium is declining over time as improvements in information and the growth of markets reduce investors' aversion to equity.

A number of forms of market failure have been considered as explanations of the equity premium. First, there is the absence of markets in which individuals can insure themselves against systematic risk in labour income and noncorporate profits. Mankiw (1986) and Weil (1992) argue that the presence of this uninsurable background risk may reduce willingness to hold additional systematic risk in the form of equity and thereby account for an equity premium. Mankiw stresses the point that, *ex post*, systematic risk is concentrated on the relatively small group in the population who incur unemployment or business failure. If individuals dislike negatively skewed returns, the equilibrium risk premium will be higher when *ex post* risk is concentrated on a subset of the population.

The second set of explanations focuses on credit constraints and, more generally, on transactions costs associated with borrowing. Constantinides, Donaldson and Mehra (1998) present a model in which the young, who would otherwise issue debt and purchase equity, are constrained from borrowing. At given asset prices, this depresses demand for equity and raises the net supply of bonds. Hence, a large equity premium is required to restore equilibrium.

4.4 Implications for the socially optimal price of systematic risk

Under the CCAPM model, as with CAPM, the value of a privately-owned asset yielding returns that are uncorrelated with national income should be equal to the net present value of returns evaluated at the riskless rate of discount.

More complex issues arise in cases where the covariance of project returns with aggregate consumption is small but positive. Consider, for example, a typical transport infrastructure project. Since changes in aggregate demand will be reflected in changes in the demand for transport services, the demand for the services of such a project will vary to some extent with aggregate consumption. However, this variation will typically be small in relation to other sources of variation in traffic flows. Hence, this risk will be relevant only if the socially optimal price of systematic risk is considerably higher than implied by CCAPM.

The issue is most simply expressed in terms of rates of return. The rate of return typically demanded by private investors in a project with 'average' risk characteristic is, in general, substantially greater than the real bond rate. Conversely, the market value of a project with 'average' risk characteristics is considerably less than the expected NPV of returns discounted at the real bond rate, even after appropriate account is taken of downside risk. Viewed in this light, it is evident that the determination of the socially optimal price of systematic risk is closely related to the analysis of the equity premium puzzle.

Grant and Quiggin (2002) show that, if adverse selection problems prevent insurance against systematic risk, as in Mankiw (1986), the optimal rate of return for public projects will be less than the market rate for projects with similar risk characteristics. Grant and Quiggin consider the case where the public sector has available a menu of projects arranged in decreasing order of attractiveness relative to the market portfolio. They show that, for the marginal project, the rate of return will be greater than the riskless bond rate but less than the private-sector rate of return.

A similar result arises if the equity premium arises from characteristics of equity not incorporated in the standard CCAPM model. This point is most evident in the model of Swan (2000) where the equity premium arises from transactions costs associated with trade in imperfectly liquid equity. In this model, asset prices are determined optimally and the real bond rate is equal to the social opportunity cost of capital. By virtue of its superior ability to issue a liquid security the government enjoys a cost advantage relative to issuers of private equity. Hence, the appropriate rate of discount for public projects is the bond rate.

Use of the market price of risk would be consistent with efficiency if the observed market premium for systematic risk was an accurate reflection of the preferences of individuals trading in an efficient capital market where all risks were perfectly diversified, or if any deviations from market efficiency were small and unimportant. However, no satisfactory explanation of the equity premium puzzle consistent with approximate market efficiency has been proposed.

4.5 Case for disregarding pure risk in public projects

The balance of evidence on the equity premium puzzle supports the view that the anomalously high price of systematic risk observed in markets for corporate equity arises because markets fail to spread risk in the manner required by the efficient capital markets hypothesis.

It is important to emphasise once again that these arguments apply only to pure risk, or, more precisely to pure systematic risk. In evaluating the project, it is necessary to take full account of ‘downside risk’ and to ensure that the cash flows used in estimation are as close as possible to the mean values, rather than being modal or ‘no surprises’ estimates. Similarly, in assessing the cost of capital, it is necessary to take account of any government guarantees to borrowers, and to value these at their expected cost. The expected cost is equal to the value of the guarantee multiplied by the probability that it will be exercised.

4.6 Public and private projects

Over the last two decades, an increasing proportion of public projects have relied, in some form, on private investment. In all such cases, it is necessary to consider the issues of risk and discounting bearing in mind that different parties may face different discount rates and different costs of dealing with risk from various sources. The appropriate project design will take account of the need for appropriate risk allocation.

For example, it is generally agreed that, where possible, the firm engaged in construction of a project should bear risk associated with construction, for example by paying penalties for late completion and receiving bonuses for early completion. This view reflects the more general idea that, in the presence of asymmetric information, the better-informed party should bear more of the risk. More generally, whenever mechanisms for risk transfer are not perfectly efficient, the allocation of risk will be relevant in the evaluation and design of projects. This is the converse of Hirshleifer's (1965) observation that, in the presence of efficient markets, the cost of risk is the same whoever bears it, since all agents face the same state-contingent prices.

The general principle that risk should be allocated to the party best able to bear it is now widely accepted in discussion of partnership and contractual relationships between the public and private sectors (Victorian Treasury 2001). However, the implications of the principle of effecticent risk allocation have not been fully developed in relation to asset ownership. Much discussion of these issues is based on an implicit neutrality proposition, namely, that the benefits associated with an investment are the same whoever bears the risks of ownership.

In general, the owner of an asset may be defined to be the recipient of the residual income generated by that asset after payments have been made to input suppliers, including lenders and bondholders. The owner of an asset is exposed to non-diversifiable risk in demand for the services generated by the asset and to risk associated with discretionary decisions about the use of the asset.

These risks can be reduced through market transactions and contractual arrangements. For example, non-systematic risks may be diversified through insurance or futures markets. Similarly, if choices about the use of assets can be anticipated in advance, contractual provisions can be used to spread the resulting risk. However, it is not, in general, possible to eliminate the risks associated with ownership or to separate them from the systematic component of demand risk.

Separation of ownership risk from the systematic component of demand risk would, of course, be possible in the presence of complete state-contingent markets, With complete markets, all risks can be traded freely and the risk allocation problem vanishes.

Separation would also be possible if all systematic risks were freely tradable. Some systematic risks, such as those affecting the profits of corporations are tradable, but others, such as the risks associated with unemployment for individuals, are not. The success of the CAPM in modelling the relative prices of securities embodying different levels of systematic risk has given rise to the mistaken assumption that the absence of markets for the systematic risk facing individuals is unimportant and therefore that the value of an investment project is independent of its ownership. The failure of the full CCAPM model to predict the market price of systematic risk indicates that this assumption is unsound

The value of a project will, in general, be different if it is publicly owned than if it is privately owned. The value will also be different for different forms of private ownership. For example, some assets may have greater value as the property of individuals or partnerships and others as widely held corporations. The cost of systematic risk will, in general, be greater for assets owned by individuals who are themselves exposed to undiversifiable systematic risk than for those of publicly traded corporations, and greater for assets owned corporations dependent on private equity than for publicly-owned assets funded by public debt and taxation.

Acceptance of the proposition that the price of risk for public projects should be lower than the price demanded by investors in private equity does not imply a general preference for public ownership of assets. The cost of risk must be taken into account, but so must the need for an appropriate allocation of other components of ownership risk.

In some cases, these considerations are likely to reinforce one another. For example, in the case of a road project that makes up part of a larger publicly-owned road network, the public owners of the network will make decisions that affect returns to the project in question, for example, regarding connecting roads, alternative routes and so on. Hence, as noted by the Economic Planning Advisory Commission (1995a,b) and the House of Representatives Standing Committee on Communications Transport and Microeconomic Reform (1997), private ownership of individual toll roads is unlikely to be lead to an optimal allocation of risk. In other cases, however, considerations of ownership risk will favour private ownership of the relevant capital.

4.7 Flexibility and project design

In the analysis presented above, attention has been focused on a single decision—whether or not to proceed with a proposed project at a fixed point in time. In reality, a range of alternatives is usually available. First, in many cases, a proposed project with a given design is only one of a number of alternatives

Second, it may be possible to vary the timing of a project. In particular, it is usually possible to defer the implementation of a given project for, say, a year, and then re-evaluate the proposal.

A crucial concept in considering alternative approaches to the implementation of a given project is that of flexibility. Broadly speaking a plan is more flexible if it allows a greater capacity to respond to new information as it becomes available. In most cases, flexibility is enhanced the later any given decision affecting the flow of returns from the project must be made.

A full-scale discussion of this issue is beyond the scope of the present paper. However, one important implication arises in relation to project evaluation in the presence of uncertainty. In considering alternative designs for projects, it is necessary to take appropriate account of the option value associated with designs that allow for flexible responses to new information. As with downside risk, the most appropriate procedure for option value is to evaluate it explicitly. Graham (1992) notes that, in evaluating projects involving option value, it is normally appropriate to use the riskless discount rate.

5 Conclusions

The problem of evaluating projects with risky outcomes is complex. Using the state-contingent approach, however, such projects can be evaluated using the same approach that has long been applied to the discounting of returns over time.

The state-contingent approach shows that discussion of risk frequently involves a confusion between two distinct concepts, pure risk and downside risk. Downside risk arises when the parameters used in project evaluation are more favourable than the relevant mean values. Pure risk is relevant when variations about the mean value are correlated with total consumption.

The analysis yields two main conclusions. First, in most cases, downside risk is more important than pure risk. Second, particularly in the case of downside risk, it is inappropriate to deal with risk by applying an adjusted discounted rate to a projection based on median or modal values. The best procedure is to represent uncertainty explicitly in order to generate unbiased estimates of mean value. If project evaluation is based on estimates of mean values, the appropriate rate of discount for public projects is, in most cases, close to the real rate of interest on government bonds.

References

Allais, M. 1953, 'Le comportement de l'homme rationnel devant le risque: critique des axioms et postulats de l'école Americaine', Econometrica 21(4), 503–46.

Arrow, K. and Lind, R. 1970, 'Uncertainty and the evaluation of public investment decisions', *American Economic Review* 60(2), 364–78.

Bureau of Transport Economics 1999, *Facts and Furphies in Benefit-Cost Analysis: Transport*, BTE Report 100, Canberra.

Campbell, H. 1997, 'Deadweight loss and the cost of public funds in Australia', *Agenda*, 231–6.

Campbell, H. and Bond, K. 1997, 'The cost of public funds in Australia', *Economic Record* 73(220), 22–34.

Chew, S. H. 1983, 'A generalisation of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox', *Econometrica* 51(4), 1065–92.

Constantinides, G. M., Donaldson, J. B., and Mehra, R. 1998, *'Junior Can't Borrow: A New Perspective on the Equity Premium Puzzle'*, Working Paper No. 6617, NBER.

Constantinides, G. M. 1990, 'Habit formation: a resolution of the equity premium puzzle', *Journal of Political Economy* 98(3), 519–43.

Department of Finance 1991, *Handbook of Cost-Benefit Analysis,* AGPS, Canberra.

Dixit, A. and Pindyck, R. 1994, *Investment Under Uncertainty,* Princeton University Press, Princeton .

Economic Planning Advisory Commission (EPAC) 1995, *Final Report of the Private Infrastructure Task Force*, AGPS, Canberra.

Economic Planning Advisory Commission (EPAC) 1995, *Interim Report of the Private Infrastructure Task Force*, AGPS, Canberra.

Feldstein, M. 1992, 'The inadequacy of weighted discount rates', 245–69 in Layard, R. (ed.), *Cost-Benefit Analysis,* Penguin, Harmondsworth.

Fisher, I. 1930, *The Theory of Interest,* Yale University Press, New Haven.

Glassman, J. and Hassett, K. 1999, *Dow 36000,* Times Books, New York.

Graham, D. 1992, 'Public expenditure under uncertainty: the net benefit criteria', *American Economic Review* 82(44), 822–46.

Grant, S. and Quiggin, J. 2002, 'Public investment and the risk premium for equity', *Economica* forthcoming.

Heaton, J. and Lucas, D. 1996, 'Evaluating the effects of incomplete markets on risk sharing and asset pricing', *Journal of Political Economy* 104(3), 443–487.

Hirshleifer, J. 1965, 'Investment decisions under uncertainty: choice-theoretic approaches', *Quarterly Journal of Economics* 79(4), 509–36.

House of Representatives Standing Committee on Communications Transport and Microeconomic Reform 1997, *Planning not Patching: An Inquiry Into Federal Road Funding,* The Parliament of the Commonwealth of Australia, AGPS, Canberra.

Kahneman, D. and Tversky, A. 1979, 'Prospect theory: an analysis of decision under risk', *Econometrica* 47(2), 263–91.

LaValle, I. and Wapman, K. 1986, 'Rolling back decision trees requires the independence axiom!', *Management Science* 32(3), 382–5.

Little, I. and Mirrlees, J. 1974, *Project Appraisal and Planning for Developing Countries,* Heinemann, London.

Loewenstein, G. and Thaler, R. 1989, 'Anomalies in intertemporal choice', *Journal of Economic Perspectives* 3(4), 181–193.

Machina, M. 1982, ''Expected Utility' analysis without the independence axiom', *Econometrica* 50(2), 277–323.

Mankiw, N. G. 1986, 'The equity premium and the concentration of aggregate shocks', *Journal of Financial Economics* 17, 211–19.

Marglin, S. 1963, 'The social rate of discount and the optimal rate of investment', *Quarterly Journal of Economics* 77(1), 95–111.

Marglin, S. 1967, *Public Investment Criteria,* MacMillan, London.

Mehra, R. and Prescott, E. C. 1985, 'The equity premium: a puzzle', *Journal of Monetary Economics* 15(2), 145–61.

Pohl, G. and Mihaljek, D. 1992, 'Project evaluation and uncertainty in practice: A statistical analysis of rate-of-return divergences of 1015 World Bank projects', *World Bank Economic Review* 6(2), 255–78.

Quiggin, J. 1982, 'A theory of anticipated utility', *Journal of Economic Behavior and Organization* 3(4), 323–43.

Sugden, R. and Williams, A. 1985, *The Principles of Practical Cost-Benefit Analysis,* Oxford University Press, Oxford.

Swan, P. 2000, ‘Does 'illiquidity' rather than 'risk aversion' explain the equity premium puzzle? The value of endogenous market trading’, No. 18 December, School of Business, University of Sydney.

von Neumann, J. and Morgenstern, O. 1944, *Theory of Games and Economic Behavior,* Princeton University Press.

Weil, P. 1989, ‘The equity premium puzzle and the risk-free rate puzzle’, *Journal of Monetary Economics* 24, 401–21.

Recent publications

REPORTS

Report 111 . . . . General Aviation: an industry overview (May-05)

Report 109 . . . . Rail Infrastructure Pricing: Principles and Practice (Jul-03)

Report 108 . . . . Rail Accident Costs in Australia (Jan-03)

INFORMATION PAPERS

Information Paper 53 . . . Australian Sea Freight, 2002–2003 (Mar-05)

Information Paper 51 . . . Focus on Regions No. 2: Education, Skills and Qualifications (Feb-04)

Information Paper 50 . . . Australian Sea Freight 2001–2002 (Jan-04)

INFORMATION SHEETS

Information Sheet 23 . . . Public road-related expenditure and revenue in Australia 2004 (Jul-04)

Information Sheet 22 . . . Freight between Australian Cities 1972–2001 (Sep-03)

Information Sheet 21 . . . Greenhouse Gas Emissions to 2020: Projected Trends for Australian Transport (Feb-03)

WORKING PAPERS

Working Paper 62 . . . . . . Modelling Responses of Urban Freight Patterns to Greenhouse Gas Abatement Scenarios (Nov-04)

Working Paper 61 . . . . . . Is the world running out of oil? An overview of the debate (May-05)

Working Paper 60 . . . . . . An overview of the Australian Road Freight Transport Industry (Dec-03)

Working Paper 59 . . . . . . Road Speed Limits: Economic effects of allowing more flexibility (Sep-03)

Working Paper 58 . . . . . . Investment Trends in the Lower Murray-Darling Basin (Oct-03)

OTHER BTRE PUBLICATIONS: http://www.btre.gov.au/publications.aspx
BTRE HOME PAGE: http://www.btre.gov.au/index.aspx
To order free publications write to:
BTRE, GPO Box 501, Canberra ACT 2601, Australia.
telephone +61 2 6274 7210, facsimile +61 2 62746816
or email btre@dotars.gov.au