



Centre For Health Economics



**Cost, Context and Decisions in
Health Economics and
Cost-Effectiveness Analysis**

Anthony J Culyer

CHE Research Paper 154

Cost, context and decisions in Health Economics and cost-effectiveness analysis

Anthony J Culyer

Centre for Health Economics, University of York, UK

June 2018

Background to series

CHE Discussion Papers (DPs) began publication in 1983 as a means of making current research material more widely available to health economists and other potential users. So as to speed up the dissemination process, papers were originally published by CHE and distributed by post to a worldwide readership.

The CHE Research Paper series takes over that function and provides access to current research output via web-based publication, although hard copy will continue to be available (but subject to charge).

Acknowledgements

I am grateful for discussions with and comments from Kalipso Chalkidou, Karl Claxton, Richard Cookson, Mike Drummond, Ali McGuire, Alec Morton, Mark Pauly, Platonas Yfantopoulos.

No ethical approval was needed.

Further copies

Only the latest electronic copy of our reports should be cited. Copies of this paper are freely available to download from the CHE website www.york.ac.uk/che/publications/. Access to downloaded material is provided on the understanding that it is intended for personal use. Copies of downloaded papers may be distributed to third parties subject to the proviso that the CHE publication source is properly acknowledged and that such distribution is not subject to any payment.

Printed copies are available on request at a charge of £5.00 per copy. Please contact the CHE Publications Office, email che-pub@york.ac.uk, telephone 01904 321405 for further details.

Centre for Health Economics
Alcuin College
University of York
York,
YO10 5DD, UK
www.york.ac.uk/che

©Anthony J Culyer

Abstract

Cost in health economics is necessarily associated with a decision. It varies according to the context of that decision: whether about inputs or outputs, the alternatives, its timing, the nature of the commitment to following a decision, who the decision maker is, and the constraints and discretion limiting or liberating the decision maker. Distinctions between short/long runs and between fixed/variable inputs are matters of choice, not technology, and are similarly context-dependent. Costs are not harms or negative consequences. Whether 'clinically unrelated' future costs and benefits should be counted in current decisions also depends on context. The costs of entire health programmes are context-dependent, relating to planned rates of activity, volumes and timings. The implications for the methods of CEA and HTA are different in the contexts of low- and middle-income countries compared with high-income countries, and further differ contextually according to budget constraints (fixed or variable).

Key words

Opportunity cost, choice, decisions, context, short/long runs, fixed/variable costs, unrelated costs, LMICs.

Highlights

Opportunity cost is harder to pin down than is commonly thought.

Some costs are inherently subjective.

Cost is not the same as harm or negative consequences.

Cost is always context-dependent.

Distinctions between short and long runs and between fixed and variable costs are chosen, not given technologically.

'Unrelated costs' should not always be included in cost-effectiveness analyses.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Cost, context and decisions in Health Economics and cost-effectiveness analysis

“Obscurities, ambiguities, and errors exist in cost and supply analysis despite, or because of, the immense literature on the subject” (Alchian 1977, 273). What Alchian could write in 1959 is apposite for today’s health economics and especially in the economic evaluation of health care technologies. The idea of opportunity cost is ill-understood (or, better, differently understood) not only by people who are not economists, but even by economists including, we may presume, health economists.

In 2005, Ferraro and Taylor (2005) put a question based on one by Frank and Bernanke (2001), to 199 economists. I have adapted it for health economists¹.

You have a health condition that could be treated without charge by your public national health insurance programme. You could also ‘go private’ by registering with a local surgical clinic offering medical care of an equivalent quality for which you would pay out of pocket, which is your next-best alternative choice. The price of private treatment is \$400 and the maximum you would be willing to pay for private care is \$500. Assume there are no other costs of either option. Based on this information, what is the opportunity cost of choosing public health care:

- A. \$0
- B. \$100
- C. \$400
- D. \$500?

The most popular answer was \$500, with 27.6% of respondents choosing this answer. The second most popular answer was \$400, with 25.6% of respondents choosing it. The third was \$0, with 25.1% of respondents choosing it. \$100 was the least popular, with only 21.6% of respondents choosing it. The answers from these well-trained economists seemed to be randomly distributed across the four possible answers. What is your answer? What is the correct answer?

According to Ferraro and Taylor, \$100 was the right answer: the value to you of the forgone care (\$500) less the \$400 out-of-pocket price you also forgo. In this case, it is essentially the consumer’s surplus forgone. But you could have taken a different view: you could have valued the forgone opportunity at its market price (\$400) which one may suppose represents the minimum sum the providers must receive for their services. Or you might have taken the view that the opportunity cost is best considered descriptively: what was forgone was private care with a bundle of associated characteristics which you say you value at \$500. Or you might say the opportunity cost was \$0, since whatever you decide, the service will be provided to someone with no additional resource costs than if it were provided to you.

This kind of confusion is plainly a serious matter, since opportunity cost is one of the half dozen key economic concepts whose understanding defines a professional health economist. One reason for confusion in this case may have been the unfamiliar context: the test was not couched in terms of the opportunity cost of resources used in the production of a good or service but rather in the

¹ The test question in the original was: Please Circle the Best Answer to the Following Question:

You won a free ticket to see an Eric Clapton concert (which has no resale value). Bob Dylan is performing on the same night and is your next-best alternative activity. Tickets to see Dylan cost \$40. On any given day, you would be willing to pay up to \$50 to see Dylan. Assume there are no other costs of seeing either performer. Based on this information, what is the opportunity cost of seeing Eric Clapton?

- A. \$0
- B. \$10
- C. \$40
- D. \$50

context of choice between a final good or service (see Parkin 2016 for other explanations). The \$100 opportunity cost is subjective and is revealed neither by the prices of the options nor in any system of accounting. The other two interpretations (C and D) follow straightforwardly: the market price of the private care (\$400) on the one hand and the general characteristics of the forgone option (choice of physician, early admission, etc.) valued at \$500 on the other.

From a public decision maker's point of view the most useful approach may be any one of the following:

- (1) In order to understand patient behaviour and the (dis)incentives required to change it;
- (2) To appreciate what must be forgone in the health care system for an expenditure of \$400 per patient on a public service contracted to a private service supplier;
- (3) To discover whether a publicly provided service package might be provided at a lower cost than a contracted-out service having somewhat different characteristics;
- (4) To dismiss the whole issue as a question of transfers rather than costs.

All four answers may therefore have something to be said for them as a reasonable interpretation according to context. What is reasonable seems to depend on context: what is the question, who is the decision maker and what is the decision maker seeking to achieve? It should not be surprising that opportunity cost will vary according to identity (e.g. position in a hierarchy of decision makers) and the objectives of the decision maker (utility maximising, profit maximising, health maximising?).

This paper is an attempt to demystify and clarify the idea of cost in economics, not only for health economists but also for analysts from other disciplines and for members of multidisciplinary teams working on economic evaluations in the field of health, health care and social care. This basic idea, that costs (and, for that matter, benefits) cannot be defined without specifying the context of the decision, turns out to be a recurring theme. I shall assume, to avoid residual ambiguity, that the cost with which we are concerned in decisions about health care investments (using, for example, HTA) are the costs of resources and that the opportunities forgone are those arising from the varied uses to which those resources might be put.

Opportunity cost and the production of services

Early emphasis on the importance of opportunity cost in economics is particularly associated with the Austrian School of economics in the late nineteenth century:

To say that any kind of production involves cost simply implies that the economic means of production, which could doubtless have been usefully employed in other directions, are either used up in it, or are suspended in it. (von Wieser 1893, 175)

This is what Robbins (1934, 2) described as Wieser's Law: "Cost' in economics is a term commonly applied to resources, most commonly 'factors of production', that have alternative uses rather than resources interpreted as items of final consumption (arguments of utility functions).' The term 'opportunity cost' may refer to the value of the most preferred alternative use to the current or to a planned use; or it may refer to the quantities of a good or service that is forgone, like an alternative drug, or an alternative nurse; or it may refer to states of being – most notably for health economists, health itself, so that the opportunity cost of a decision becomes the health that would have been generated had the decision gone in favour of the preferred alternative. Whichever of these it may be, cost is always an opportunity that is necessarily forsaken when a decision is made. In a well-informed market with low transaction costs, the value of the forgone opportunity (say, use of a resource) may be objectively revealed as a dollar amount. This represents the highest value

someone else places upon it as revealed in offers they make, and where they will need to be outbid if several are bidding, or compensated if they already own it and are to make it available to a bidder in a Walrasian auction. Within an organisation like a hospital or an ambulance service, resources also have alternative uses, some of which are within the organisation (for example, rooms can be reallocated to other uses) and some of which lie outside (for example, ambulances can be sold to other users). The external uses may reveal opportunity cost as a monetary price. The internal opportunity cost is not revealed in a market price but is subjective, existing only in the mind of the decision maker (Buchanan 1969, Wiseman 1989). This is not to say that the subjective judgment that a particular resource is best used in a particular way is not a communicable judgment. Plainly that is not the case, as managers, who often have to act upon their subjective valuations of opportunity cost, can be held to account for their decisions through audit trails, direct enquiries and so on and can give their reasons for making the decisions they made when called to account.

Cost is always the consequence of a decision and is in turn a consequence of the resource scarcity that standard texts tell us (at least since Robbins 1932) lies at the heart of economics as an analytical discipline. All economists seem to agree that the forgoing of a want or need is fundamental and not itself contextual. It is what *defines* cost. All else, however, is contextual. How the forgone option is best measured or valued will therefore depend on who the decision maker is and the question the decision maker wants answered.

Costs vs undesired attributes

A significant implication of the idea of opportunity cost is that negative consequences are not the same as costs. There is a tendency, especially in applying techniques like multi-criteria decision-analysis (MCDA), to create categories (often overlapping) of benefits and costs treated broadly as 'pros' and 'cons' (Campillo-Artero et al. 2018). Loose usage easily transforms 'harm' or 'negative consequence' into 'cost'. Consider the homely example given by Alchian, whose extensive reworking of the cost theories of the Austrian School (notably von Wieser 1893) and the early English marginalists (Wicksteed 1910, Jevons 1871, Marshall 1871) is drawn on extensively here.

Suppose you want to build a pool in your garden. "The toil and trouble of digging it and the nuisance of noisy disobedient neighbourhood children and uninvited guests who use it are undesirable attributes of the pool. They are not the costs of creating and having the pool. This distinction between (a) undesirable attributes inherent in some event and (b) the highest-value forsaken option necessary to realize the event is fundamental, for only the latter is cost as the term is used in economics" (p. 301-2). Blood, sweat and tears associated with building the pool are not costs. They may be relevant to the decision in two ways: they reduce the value of creating and having the pool and they may be, indeed probably are, attributes of some of the other human inputs in the pool-building project which may need compensation if these resources are to be made available and thereby enter into costs. But they are not costs of building the pool because they are not resources having alternative uses. If, however, you are doing the digging, then of course your time is a resource with alternative uses and is accordingly a cost.

Cost and decisions

Cost is an opportunity for using a resource that is necessarily forsaken when a decision is made. The character of the decision in question will normally affect its cost. This implies (as we have already seen) that what is considered to be a cost is context-dependent.

As has again been seen, some costs are subjective, not readily revealed in systems of accounts, and only indirectly revealed by decisions. One aspect of context relates to the decision and the nature of

the commitment being entered into. One may distinguish, for example, between a decision to acquire (acquisition cost) and a decision to operate a programme of costs and outcomes over a period of time. Of course, one may combine them if the decision is both to acquire a resource and to use it in particular ways for a period. But one may also consider these as two (successive) decisions. Since the period over which the production programme is to be run is also a matter of choice, there is an opportunity cost of each. The cost of acquisition is a decision to buy. Since it may not be a decision to also hold the acquired asset, the actual cost of buying is simply the immediate depreciation in value when the resource changes ownership: the difference between the purchase price and its immediate resale value. The decision to operate a programme also entails expenditures (usually called variable costs) that are incurred over time and whose present (discounted) value is the cost of running the programme at the time of decision. The variable costs in health care, as elsewhere, are typically a function of the rate of activity, the planned volume of work, the date at which it is to begin, and the length of the period over which it is to run.

A typical health care programme might embody both the acquisition and the continuing use of a resource, usually combined with others, over a planned period, with planned or anticipated rates of utilisation by a target population. These are all matters for the decision maker to decide and different decisions (according to context) generate different costs and benefits. In HTA or CEA it is usual to consider first the inherent potential of the resource viewed as an 'intervention' or 'technology'. One category of consideration is the expected direct health benefit per individual and any associated positive or negative health-related consequences as viewed by the decision maker. These would typically be discounted over the lifetime of the individual beneficiaries. The opportunity costs would be discounted over the period in which costs are incurred (the maximum being the remaining lifetime of each beneficiary). A second category to consider is the intended population who will receive the intervention. The present value of the programme costs of the decision to provide a technology of a specified type for a specific population under specific circumstances for a specified period would then constitute the general opportunity cost of the decision to run the programme. That is, the opportunity cost of the resources in their most preferred alternative use – wherever that might be.

The present value of a cost or benefit accruing in the distant future is quite small. For example, at 3%, a dollar in 50 years' time has a present value of just above 2 cents and over 30 years of about four cents. At 8% the present values are one cent in 30 years' time and one fifth of a cent after 50 years. The importance of accuracy in distant costs and benefits is evidently less than that of closer costs and benefits.

Even a high-level health care decision maker is typically required to operate within a budget. This is as true of a private sector insurer as it is of a public sector insurer. How they optimise their decisions is likely, however, to depend on different criteria. A public sector payer might typically specify one criterion (among others) such as maximising the impact of the available budget on population health. In such a case, the opportunity cost of investing in one programme rather than another may be most usefully considered in terms of the forgone population health that the most preferred alternative use would have been expected to have generated. This is not subjective and can be empirically estimated (as by Claxton et al. 2012).

The determination of the budget is a matter for decision makers higher in the hierarchy of decision; for example, the senior management team and board of a commercial insurer, the senior management and trustees of a non-profit payer or a cabinet of ministers in a government. In each case, the decision maker is appointed by and is accountable to a still higher authority, which is the ultimate source for the values of the organisation as stated in, for example, its vision, aims and objectives, and in any specific obligations, some of which may be legal requirements, placed upon

the decision maker. The governance arrangements and lines of authority and accountability of a decision-making organisation constrain decision makers and, by the same token, constrain the role of economic advisers, whose accountability is typically to the decision maker. Analysts are, of course, at liberty to state their own values and erect economic evaluations accordingly, without any accountability other than to their own consciences, but such studies would be a species of advocacy unless there were a coincidental symmetry between these values and those of decision makers in the relevant sector.

Scope of decisions

The programme scope (whom covered, for what conditions and over what period) is a matter of choice for the decision maker. The scope of a decision defines the perspective of an economic evaluation and therefore the alternative uses of resources to be considered. The scope essentially specifies the matters of interest to the decision maker and generally includes such considerations as the following: the technology or technologies to be evaluated and their comparators; the character of benefits to be assessed; the expenditures (budget impact) and costs to be considered; the types of individual whose interests as patients, family carers, clinical professionals, service providers, product manufacturers and so on, are to be embraced; the measures of benefit; the periods over which costs and benefits are to be measured, and any harms or non-health benefits that may accrue. These – by no means a complete list - all have the characteristic that they require social value judgements to be made about their inclusion or exclusion and the weight to be attached to some of them when they are combined. In addition, there are technical specifications which do not embody social value judgements as distinct from judgements, some of which may be value judgments, concerning scientific method. These judgments about 'good science' are sometimes embodied in a reference case (e.g. Wilkinson et al. 2016) requiring social virtues like explicitness about the value judgements used, transparent evidence and modelling that enables replication by other scholars, and other scientific values like appraising the quality of evidence (robustness, precision, etc.) and the use of efficient sensitivity analysis.

While defining scope and making social value judgments are the responsibility of decision makers, it does not follow that analysts have no role to play. On the contrary, they are by experience better at distinguishing between social value judgments, other kinds of value judgment (e.g. as to what is 'good' science), and other kinds of judgment, both qualitative and quantitative. They can brief decision makers on the ethical implications of possible decisions, and they can seek logical consistency. Decision makers are apt to make some decisions for illegitimate reasons, or without weighing the alternatives. For example, in deciding the scope of a specific HTA a decision maker might implicitly exclude some consequences through oversight, for political convenience, to avoid inconvenient conflicts, or through inability to imagine some of the collateral possibilities. The analyst should be able to alert decision makers to such possibilities so that the decision, when taken, can be seen to be a considered outcome. This does not require the analyst to usurp the decision maker's accountable role, but rather to support it in principled ways.

Scope is also contextual. Consider the perspective adopted in a low- or middle- income country by an internal decision maker in the Ministry of Health compared to that adopted by an external aid giver. At the aggregate level concerning the size, target population and periodicity of the aid, the opportunity cost as seen by the aid giver might be seen exclusively in terms of alternative countries, populations and dates while that seen by the internal decision maker will characteristically concern the consequences of substitution effects. An example of substitution would be if domestic funding is replaced in part or whole by aid funding and domestic resources are released for other uses (including non-health uses like better roads and more secondary education). Internally, the opportunity cost of declining the aid includes the loss of a variety of health and non-health benefits

of which only some may be of concern to the aid giver. The opportunity cost for each party is defined by its own objective function and budget constraint.

Related and unrelated costs

Morton et al. (2016) briefly reviewed a long-standing controversy over the treatment of 'unrelated' costs (and benefits), defined as consequences of a decision that prolongs life and hence may involve future costs (and benefits) that would not otherwise have been incurred. They make a strong argument for a balanced approach, that is, either include unrelated costs and benefits or exclude both. They argue, following several other contributions that the correct approach is to include both (suitably adjusted, of course, for timing and uncertainty). Is this correct?

The first notable problem with the recommendation of Morton et al. is that it is vulnerable to an *ad nauseam* problem of infinite regress. It is one thing to consider the costs of future expected treatments that result from extensions to life but another to take account of the further life extensions that these future interventions may make possible, the new children who may be born who would not otherwise have been born, and who will doubtless have future health care needs. Indeed, if *all* unrelated costs are, as a matter of principle, to be considered, no matter how remote in time or geography, then there appears to be no limit to the costs to be considered other than those imposed by discounting and time preference on the one hand and the sheer cost of making the necessary estimates on the other. There is, however, a deeper problem with including unrelated costs (and benefits). This arises from their interpretation of official guidance on the conduct of economic appraisal in the public sector (HM Treasury 2003) as follows: "the test for including or excluding a cost or benefit depends on whether it flows from the commitment which we are being asked to make at the moment of decision." Indeed that is so. However, we need to be satisfied that a consequence does indeed flow from the decision and to enquire as to the nature of that 'commitment'.

The idea of 'commitment' implies some form of contract, for example one having the nature that an extended survival consequent to the current decision will entitle an individual to such further (clinically unrelated) interventions as may be deemed needed (or demanded). Such contracts cannot be taken for granted. In a health care system having a defined benefits package, future unrelated costs of interventions not in the package at the future date will not be incurred by the system's current decision maker and the present value of any expected expenditures made by others are not costs of the current decision. By contrast, in an established system that is comprehensive in its coverage of people and treatments, the reasonable expectation is that there is a commitment to meet future consequential but unrelated costs, including those to be incurred by the as yet unborn descendants of the present generation of beneficiaries.

It thus becomes clear that the treatment of unrelated costs is context-dependent. In Low- and Middle- Income Countries (LMICs) only now venturing into long term strategy for universal health coverage, if the reasonable expectation is that future unrelated care needs will not be covered in the benefits package, then they should not be included as costs of current decisions, unlike the future costs of treating clinically related conditions (unless these are exempted by the decision maker as an explicit social value judgment). In countries where the expectation is that unrelated treatments will be delivered, or in any scheme in which future unrelated consequential care is provided for, then unrelated costs ought to be included in the costs of the current decision. In private health insurance schemes, whether or not unrelated future costs are properly included depends again on the contract.

Benefits and costs should in general be treated symmetrically. If, for example, future unrelated costs are not covered by the contract, then estimates of future benefits that implicitly include the benefits of future unrelated interventions should be adjusted downwards and, in the reverse case when unrelated treatments are included, estimates of the present value of subsequent benefits ought also to be included.

Some past advocates of inclusion (e.g. Morton et al. 2016) take NICE to task when it prescribes in its NICE Guide to the Methods of Technology Appraisal (NICE, 2013) “costs that are considered to be unrelated to the condition or technology of interest should be excluded”. On the contrary, the kind of implicit contract represented by the NHS describes precisely the circumstances under which such costs ought to be taken into account. Use of life tables to estimate benefits from NICE-approved interventions almost certainly include the benefits of future unrelated interventions, which suggests further that NICE’s methodology is inconsistent in its asymmetrical treatment of costs and benefits.

Rates and volumes

There is an ambiguity in the idea of an outcome or output in health care, as elsewhere, which can mislead. The ambiguity is this: are we speaking of rates or volume of output? The relation between cost and rates or volumes in economics is in general qualitative, in the sense that the relationship can be signed as positive or negative. Whether the effect size is substantial or not is entirely an empirical matter. The following empirical generalisations may probably be deployed by health economists to good effect:

Postponed implementation of a decision has a lower cost than immediate implementation. Reasons for this include improved information enhancing the effective use of medicines and other inputs and the possibility of lower future prices through negotiation. When a decision maker wishes to implement a decision very quickly (one might say ‘in the short run’) the present value of cost will be higher than otherwise. A classic example is the response to an epidemic. The case for advance stockpiling of medicines and fast-response teams of specialists is that they are relatively low-cost means of achieving the benefits of an emergency intervention. At the time of actually using these resources, however, the alternatives forgone at the time of their creation are sunk costs and quite irrelevant to the current choice situation, when the opportunity costs relate to the alternative uses of the stocks and teams rather than to their creation.

Sunk costs should generally be ignored in decision-making, other than as learning experiences that may inform future decisions. The fact that resources were forgone in a decision that is already made is not a determinant of the current alternative use value of resources. Thus, the fact that large sums may have already been expended on a health project is not an argument for completing the project if the current assessment of the opportunity cost of completing it exceeds the present value of the benefits of completion. It may be that heavy past expenditures are deployed in a face-saving exercise, but that represents an investment in something other than health or health care.

Distinctions between ‘short’ and ‘long’ runs are context-dependent and not for arbitrary measurement. A short run is one in which the decision maker *chooses* not to vary one or more inputs, i.e. they are treated as ‘fixed’ even though they could in principle be varied. The long run is when all, or at any rate a significant number of inputs, are chosen to be varied. Neither, despite their names, has anything to do with periods of time. Both are determined by an interaction between the desires of the decision maker and the costs of changing specific inputs, some of which may be changed quickly with relatively low adjustment costs, but others not. In short, the decision maker chooses the ‘run’; it is not a technically fixed or exogenous variable.

Whether a cost is treated as fixed or variable is context-dependent. Thus, a contracted rental lease on a building or a contract of employment may be taken as 'given' in many contexts involving the use of the resources under contract. Contracts can, however, always be renegotiated (at a cost) or breached (at a cost) and so become variable when decision makers judge the circumstances to require adjusting a fixity.

Resources that are apparently technically fixed are also variable in some context. Buildings and railway track are examples of resources that seem fixed for most purposes. However, both are subject to more or less constant maintenance and minor alteration so that, over time, substantial increases or reductions in infrastructure are achieved. They can even be varied very quickly when it is so wished. When the English Great Western Railway's old broad-gauge track was changed to the modern standard narrow gauge in 1892, the entire stretch of 213 miles from Exeter to Penzance was changed in one weekend (Smith 1985). Thus, for some decisions (e.g. usual timetabling and ticketing) the track is assumed fixed and for others it is assumed variable.

Resources vary in the costliness of their variability. In the 1892 conversion of English railway track, 177 miles of the stretch of line between Exeter and Penzance had to be altered from the old longitudinal timbers to the modern cross-sleepers (ties) all in one weekend. It took an army of platelayers to do it – 4200 of them. The only truly fixed factor is either one for which the cost of varying really are infinite or one that the decision maker wants to vary literally immediately.

Time. It is useful when planning the production and delivery of health services to consider the impact on the present value of costs of:

- The starting date (postponement as already mentioned).
- The closing date (up to a possible date at which discounting renders the present value negligible).
- The rate of output (i.e. the output or outcomes expected per period).
- The volume of output (i.e. the cumulative output or outcomes between the starting and closing dates).
- The faster the rate of an activity, in general the higher the present value of its cost for any planned programme period. This is a standard expectation of cost functions, implying that marginal cost is positive. It may not hold universally, however. Stocks of medicines, for example, may not increase in proportion to the rate of a clinic's activity. Again, context matters and conventional assumptions always need checking.
- The greater the volume of activity the greater the present value of a programme's cost. The volume of a service is the cumulative output over a defined period, i.e. a planned programme. Plainly, given both the date at which a programme begins and a constant rate of activity, enlarging its volume implies a more distant endpoint. Alternatively, with a given (chosen) endpoint, volume may increase only if the rate of activity also rises.
- The greater the volume of activity given a constant rate of output, the lower the marginal cost (i.e. the present value of cost rises but at a decreasing rate). The reason for this volume effect is usually that planning for larger volumes enables the use of more cost-effective methods. For example, at low volumes patients may be referred elsewhere for some diagnostic tests, whereas a larger volume makes desirable the purchase and use of a scanner of one's own. The effect is often referred to as a 'scale economy'.
- The larger the volume, the lower the present value of cost of each unit of output/outcome (average cost).
- If both the rate and the volume of activity increase, the impact on marginal cost is ambiguous and is entirely context-dependent, i.e. requiring specific evidence from trials, pilots and the like.

Social cost, externalities and publicness

An external cost is a cost imposed on others than the decision maker. As with 'internal' costs, one should distinguish between opportunity costs and harms. An external harm occurs when an infectious person communicates a disease to others having a direct negative impact on their welfare. There may also be external opportunity costs, for example in the form of defensive health care to mitigate the transmitted disease or reduction in productivity through absenteeism and presenteeism. External costs (of interest to health economists) include the consequences of hazards created by other people's behaviour (e.g. in traffic) and pollution by industry and private households. Social costs are the sum of internal (private) and external costs. It is unhelpful to treat external costs as equivalent to social costs (as is sometimes found). External costs, when they exist, along with internal costs, are an element of social cost. When they are absent, social cost is the same as internal cost.

External costs generally arise in the absence of appropriate property rights, specifically in the ability of people to be able to establish titles, for example, to a bug- and pollution-free environment – or its converse. With tradeable rights, either the imposer of the cost must compensate those on whom the costs are thrust or, in the converse case, the polluter must be compensated for reductions in pollution (Coase 1960). Since many of these external effects have the character of publicness, there will often be a need for public regulation of tax/subsidy arrangements that, for example, subsidise vaccination programmes and establish markets in pollution rights.

The optimal level of pollution, like the optimal incidence or prevalence of a common disease, is unlikely to be zero. There are two main reasons for this. One is that the marginal benefit falls as the external cost or harm is increasingly controlled, while the marginal cost of controlling it commonly rises. The optimal level will therefore typically be higher than zero because the social value of further elimination is less than its social cost. In public health care systems with a fixed budget, exceeding this optimal level implies that more health is being sacrificed than further elimination will generate. The other reason is that there are many other opportunities for public expenditure on health, and a policy of pollution elimination to benefit health is unlikely to commend itself if that expenditure might generate a greater health benefit by being applied to other health care activity. This is an example of the opportunity cost where spending health care resources on one programme is the health lost from not spending it on another.

The economic 'publicness' of the effects of these decisions has little to do with the ownership of resources, which may be either public or private. The critical characteristic of a public good or a public bad in economics is that the effect (cost or benefit) cannot be individualised. As immunisation programmes approach herd immunity, *everyone* benefits; as road traffic increases without regulation, *all* highway users are exposed to increased risk. The choice between the use of public or private agents will rest partly on their relative efficiency, partly on their context-specific characteristics like human resources practices, partly on the costs and enforceability of contracts and governance structures, and partly on general social conventions regarding the accepted roles of the private and public sectors.

Conclusions

It turns out that cost, or opportunity cost, is not quite the simple idea so often claimed for it. Cost certainly represents a forgone opportunity but what is forgone may be a physical entity, an aspect of human flourishing, or the highest value attached to either. The customary view in economics has been that the term 'opportunity cost' relates to the alternative uses of resources. It is not necessarily captured in systems of financial accounting, even when expressed as a monetary value. It is

sometimes perceived only subjectively in, for example, decisions about the allocation of resources within an organisation. It is highly context-dependent. It is always the result of a decision but who the decision maker is and what discretion they are allowed are contextual. Likewise, what a decision commits the decision maker to may be highly variable and is again context-dependent (like the stability and comprehensiveness of a benefits package in a public health care system or National Health Service). The decision's essential characteristics (timing, duration, rate and volume) are all choice variables for the decision maker and those to whom the decision maker is accountable, all of whom define the scope of the decision and its other relevant characteristics.

The major take-home lessons are that cost in decisions ought always to be an opportunity cost and that cost is always context-dependent or contingent. It is contingent upon the decision to be taken, the identity and objectives of the decision maker, the budgetary, technical and political constraints imposed upon the decision maker, their permitted discretion within these constraints, their accountability to various stakeholders, and the judgments (value and scientific) of the decision maker as to the considerations deemed relevant in any particular context.

References

Alchian, AA (1959) Costs and outputs, in M Abramovitz et al. *The Allocation of Economic Resources*, Stanford: Stanford University Press, reprinted in AA Alchian, *Economic Forces at Work*, Liberty Press, 1977, 273-299.

Buchanan, J M, (1969) *Cost and Choice: an Inquiry in Economic Theory*, Chicago: University of Chicago Press.

Campillo-Artero, C, Puig-Junoy, J, Culyer, AJ (2018) Does MCDA trump CEA?, *Applied Health Economics and Health Policy*, <https://doi.org/10.1007/s40258-018-0373-y>.

Claxton K, Palmer S, Longworth L, Bojke L, Griffin S, McKenna C, et al. (2012) Informing a decision framework for when NICE should recommend the use of health technologies only in the context of an appropriately designed programme of evidence development. *Health Technol Assess*. 16: 1–323.

Coase, RH (1960) The problem of social cost. *Journal of Law and Economics*. 3: 1-44.

Ferraro PJ, Taylor LO (2005) Do economists recognize an opportunity cost when they see one? A dismal performance from the dismal science, *Contributions in Economic Analysis & Policy*, 4. doi:10.2202/1538-0645.1469.

Frank, RH. and Bernanke BS. (2001) *Principles of Microeconomics* New York: McGraw-Hill/Irwin.

HM Treasury 2003 [updated 2011]. The Green Book: Appraisal and Evaluation in Central Government, TSO, London. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220541/green_book_complete.pdf. Accessed 11/06/2018.

Jevons, WS (1871) *The Theory of Political Economy*. Reprint of 1931 edition, Charlottesville: Ibis.

Marshall, A (1871) *Principles of Economics*. London and New York: Macmillan.

Morton, A., Adler, A., Bel, D., Briggs, A. et al. (2016) Unrelated future costs and unrelated future benefits: reflections on NICE guide to the methods of technology appraisal, *Health Economics*, 25: 933-938.

NICE 2013. Guide to the Methods of Technology Appraisal, London, NICE. Available at <https://www.nice.org.uk/article/pmg9/chapter/foreword>. Accessed 11/06/2018.

Parkin, M (2016) Opportunity cost: A re-examination, *Journal of Economic Education*, 47:1, 12-22, DOI: 10.1080/00220485.2015.1106361.

Robbins, LC (1932) *An Essay on the Nature and Significance of Economic Science*. London and New York: Macmillan.

Robbins, LC (1934) Remarks upon certain aspects of the theory of costs, *Economic Journal*, 44: 1-18.

Smith, LA (1985) 'The broad gauge story', available online at: <http://lionels.orpheusweb.co.uk/RailSteam/GWRBroadG/BGHist.html> (accessed 23 February 2018).

von Wieser, F (1893) *Natural Value*, London and New York: Macmillan.

Wicksteed, PH (1910) *The Common Sense of Political Economy*. 1933 edition, London: Routledge and Kegan Paul

Wilkinson, T, Claxton, KP, Sculpher, MJ, Revill, P . et al. (2016) The International Decision Support Initiative reference case for economic evaluation: an aid to thought, *Value in Health*, 19: 921-928.

Wiseman, J (1989) *Cost, Choice and Political Economy*, Cheltenham: Edward Elgar.