

This is a repository copy of *Pricing implications of non-marginal budgetary impacts in health technology assessment : a conceptual model*.

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/135408/>

Version: Published Version

---

## **Monograph:**

Howdon, Daniel David Henry and Lomas, James Richard Scott  
orcid.org/0000-0002-2478-7018 (2017) Pricing implications of non-marginal budgetary impacts in health technology assessment : a conceptual model. Discussion Paper. CHE Research Paper . Centre for Health Economics, University of York , York, UK.

---

## **Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

## **Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



Centre For Health Economics



**Pricing Implications Of  
Non-Marginal Budgetary  
Impacts in Health Technology  
Assessment:  
A Conceptual Model**

Daniel Howdon, James Lomas

**CHE Research Paper 148**

# Pricing implications of non-marginal budgetary impacts in health technology assessment: a conceptual model

<sup>a</sup>Daniel Howdon

<sup>b</sup>James Lomas

<sup>a</sup>Department of Economics, Econometrics and Finance, University of Groningen, Netherlands

<sup>b</sup>Centre for Health Economics, University of York, York, UK

November 2017

### **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**

The authors thank Jochen Mierau, Mike Paulden and Simon Walker, as well as participants at the Health Economists' Study Group Summer 2017 Conference at the University of Aberdeen, for insightful comments on earlier versions of this work. Any remaining errors are our own.

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/](http://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](mailto:che-pub@york.ac.uk), telephone 01904 321405 for further details.

Centre for Health Economics  
Alcuin College  
University of York  
York, UK  
[www.york.ac.uk/che](http://www.york.ac.uk/che)

© Daniel Howdon, James Lomas

## **Abstract**

This paper introduces a framework by which to conceptualise the decision-making process in health technology assessment when the quantity of health forgone by acceptance is high enough such that the use of a single threshold based on the marginal productivity of the health care system is inappropriate, and draws out the implications of this for pharmaceutical pricing. Under the condition of perfect divisibility, a large budgetary impact of a new treatment may imply that optimal implementation may be partial rather than full, even at a given incremental cost effectiveness ratio (ICER) that would nevertheless mean the decision to accept the treatment in full would not lead to a net reduction in health. In a one-shot price setting game, this seems to give rise to horizontal equity concerns which may be more apparent than real. When the assumption of fixity of the ICER (arising from the assumed exogeneity of the manufacturer's price) is relaxed, it can be shown that the threat of partial implementation may be sufficient to give rise to an ICER at which cost the entire potential population is treated, maximising health at an increased level, and with no contravention of the horizontal equity principle.

**JEL:** I10, I11, J18.

**Keywords:** health technology assessment, economic evaluation, budgetary impact.



## 1. Introduction

Conventional decision-making for health technology assessment relies upon accepting or rejecting new interventions based upon their incremental cost-effectiveness ratio (ICER) relative to some benchmark threshold,  $\lambda$  (Drummond et al. 2015)<sup>1</sup>. More recently, following decisions made in the UK regarding drugs for the treatment of hepatitis C which, while judged to be cost-effective, according to such a benchmark threshold, by the National Institute for Health and Care Excellence but unaffordable by NHS England (National Institute for Health and Care Excellence 2015a; National Institute for Health and Care Excellence 2015b; National Institute for Health and Care Excellence 2015c), attention has become focused on a fuller consideration of the likely opportunity costs of such interventions, adjusting for the size of their budgetary impact. Given the likely diminishing marginal effectiveness of spending on healthcare in improving health and associated non-linearity of the relationship between healthcare expenditure and health outcomes, larger budgetary impacts are likely to imply larger health losses per pound spent on healthcare (McCabe et al. 2008; Harris 2016; Culyer 2016; Paulden et al. 2017). Empirical estimation of this relationship has suggested that the use of a constant threshold at the margin of the health system's expenditure is likely to lead to substantial forgone population health where the total budgetary impact is large, with the marginal threshold of £12,936 empirically estimated by Claxton et al. (2015) deemed to be around 7% too high in the case of a new treatment with a budgetary impact of £2.5bn (Lomas et al. 2017)<sup>2</sup>. The importance of using the correct threshold is clear in the light of recent cases involving new treatments with potentially high budgetary impact, such as that of sofosbuvir, one of the hepatitis C drugs recently considered for approval in the UK, which was estimated at around £772mn in a single year (National Institute for Health and Care Excellence 2015b).

The assumption of a constant threshold may be unproblematic in the case of decision-making which does not use up a large proportion of the budget of the health service, faced with healthcare producing health relatively 'flat-of-the-curve' with respect to expenditure at the margin. The use of a benchmark threshold in such a first-best world (see Section 1.1) is held to be a simplification of a mathematical programming problem which in its full form would rely on information that is not available to decision-makers, and require impractical constant funding and defunding of services, with associated costs of so doing (Epstein et al. 2007). Furthermore, in such situations, decision-making at the margin does not provoke concerns regarding horizontal equity. The comparison of a constant threshold with a constant ICER means that a treatment is either provided (where  $ICER < \lambda$ ) or not provided (where  $ICER > \lambda$ ) in its entirety for all patients with equal capacity to benefit from treatment.

Where the budgetary impact is greater, partial (but not full) acceptance may prove cost-effective, due to the non-linearity of the relationship between expenditure and health outcomes. Previously, such issues have appeared to be, or been held to be, irrelevant for two reasons. First, in the absence of an estimate of this relationship beyond, but close to, the margin, partial acceptance can be justified only with resort to a full mathematical programming model. However, in the light of recent empirical estimation of such a relationship, a reduced form model which takes account of this non-linearity is now possible without the need to resort to a full mathematical programming solution. A second objection has been that the partial funding of services is held to contravene principles of horizontal equity, where individuals in equal need (with equal capacity to benefit) should be given equal healthcare treatment (Epstein et al. 2007). This paper argues that the claim that partial provision would impose a unique contravention of the principle of horizontal equity is inaccurate, and furthermore that any apparent such problem is potentially obviated when the price of treatment is endogenous rather than assumed to be exogenously given.

This paper is structured as follows. Section 2 presents a four-quadrant graphical framework that proposes a reduced form account of decision-making when such situations arise, without resorting to a full mathematical programming solution. Section 3 discusses an allocation problem within the framework of the model, treating the ICER as exogenous. Section 4 considers the horizontal equity implications of partial provision when the ICER is treated as exogenous, and provides a framework with which to illustrate the costs in terms of population health of the imposition of a strict constraint enforcing horizontal

<sup>1</sup> Equivalently, a net benefit rule (Stinnett and Mullaly 1998) can be used. Without loss of generality, this paper will discuss decision-making solely using the method of ICER/threshold comparison

<sup>2</sup> Note that this marginal estimate itself is substantially lower than the threshold of around £30,000 used by NICE in practice

equity in acceptance only. Section 5 considers how a situation where partial acceptance is optimal in a static problem might be resolved dynamically given the optimal reactions of both the decision-maker and the manufacturer when the ICER is treated as endogenous (arising from the price chosen by the manufacturer) rather than exogenous. Section 6 concludes.

### 1.1. What does the threshold represent?

The meaning attached to the benchmark threshold requires some exposition. A comprehensive summary of the various normative and technical interpretations attached to its value, and the implications thereof, is provided by Culyer (2016). We initially consider a budget-constrained healthcare system operating on its production possibility frontier (PPF), where (as detailed in Paulden (2016)) any healthcare provided by the acceptance of a new treatment displaces (parts of) existing treatment programme-populations, with the least cost-effective (parts of) existing treatment programme-populations displaced first, thus displacing as little as health as possible for any given level of displaced spending. In such a system, the threshold we first here consider for marginal decision-making represents the value implied by the efficiency of the system at the margin, representing the opportunity cost of health displaced by a new treatment with a marginal budgetary impact or the reciprocal of the shadow price of the budget constraint (a 'first best' threshold, as characterised by Culyer (2016))<sup>3</sup>. While this is a particularly restrictive assumed model and unlikely to entirely accurately characterise anything beyond an ideal type healthcare system, it is highly useful for expository purposes. Many departures from these assumptions can be incorporated within the model, and we later highlight one relaxation of this assumption and the corresponding change in interpretation arising from it.

---

<sup>3</sup> This value is commonly termed  $k$  in existing literature when referring to this specific conception of the generic threshold  $\lambda$ . For consistency and clarity, we term this value  $\lambda_{\text{marginal}}$



## 2. Model

Assume that a health service is able to perfectly divisibly assign healthcare inputs, with any existing resources in use able to be reswitched to alternative purposes costlessly and immediately. Health, provided by the health system, is a function of inputs used for treatment, with this relationship specifying a health production function. The decision-making process faced by the relevant authority requires the maximisation of population health, producing according to this production function, subject to its budget constraint.

$$\max \left( \sum_{l=1}^L \sum_{j=1}^{J_l} \sum_{i=1}^{I_l} H_{ijl} x_{ijl} \right)$$

subject to:

$$\sum_{l=1}^L \sum_{j=1}^{J_l} \sum_{i=1}^{I_l} c_{ijl}^H x_{ijl} \leq b$$

and

$$0 \leq x_{ijl} \leq 1, \quad i = 1 \dots I_l \quad j = 1 \dots J_l, \quad l = 1 \dots L$$

where  $H$  is health produced by treatment  $j$  with cost  $c$  within programme  $l$ , for fraction  $x$  of population  $i$ , and  $b$  is the total budget available. Given that this entails selecting the most cost-effective treatment first, followed by the next most cost-effective treatment and so on, this implies a falling marginal effectiveness of treatment as the budget increases or, by the same token, as more health is produced. The solution of such a problem in its full form requires greater information than is available to, and greater flexibility than is practical for, the decision-making authority and, given these constraints, decision-making generally takes the form of comparing the ICER arising from the new intervention to a constant threshold,  $\lambda_{\text{marginal}}$ .

While the functional form of a health production function (HPF), of the type specified above, may be unknown, we can consider some general solution of this maximisation problem which implies an indirect health production function (IHPF) of the form  $H = f(b, \mathbf{c})$ , where  $\mathbf{c}$  is a vector of treatment costs<sup>4</sup>. Recent work has produced empirical estimates of the slope of such an indirect health production function beyond the threshold, taking account of its non-linearity (Lomas et al. 2017), enabling decision-making to potentially go beyond such a naive assumption. Although such empirical estimates do not recover the explicit production function drawn out above in full and the optimal mix of treatments for programme populations, the values derived by such estimation do imply an IHPF, holding individual treatment costs constant and varying  $b$ .

Assume a new healthcare intervention is now proposed, with some large budgetary impact  $m$  if fully implemented, for a condition which is currently not treated (i.e. the ICER is based on a comparison to no treatment, and the ICER is equal to the cost per QALY gain over no treatment). We initially assume constant QALY returns to spending on this intervention. Health ( $H_1$ ) can be produced by use of this new intervention, or ( $H_2$ ) by the use of existing treatments already employed. The decision-making authority is tasked with maximising overall population health ( $H_1 + H_2$ ), and decides upon whether to approve or reject the new intervention on this basis. The monetary price of the treatment, and the associated consequent ICER, is initially assumed to be given and fixed. This implies the displacement of treatment with total expenditure of up to  $m$ , depending on the fullness of implementation.

Consider Figure 1, a graphical illustration of this situation.

<sup>4</sup> In line with the use of 'indirect production function' in the conventional theory of the firm, we use the term 'indirect health production function' to distinguish this, as a function of prices and a budget, from a health production function, whose arguments are inputs explicitly.

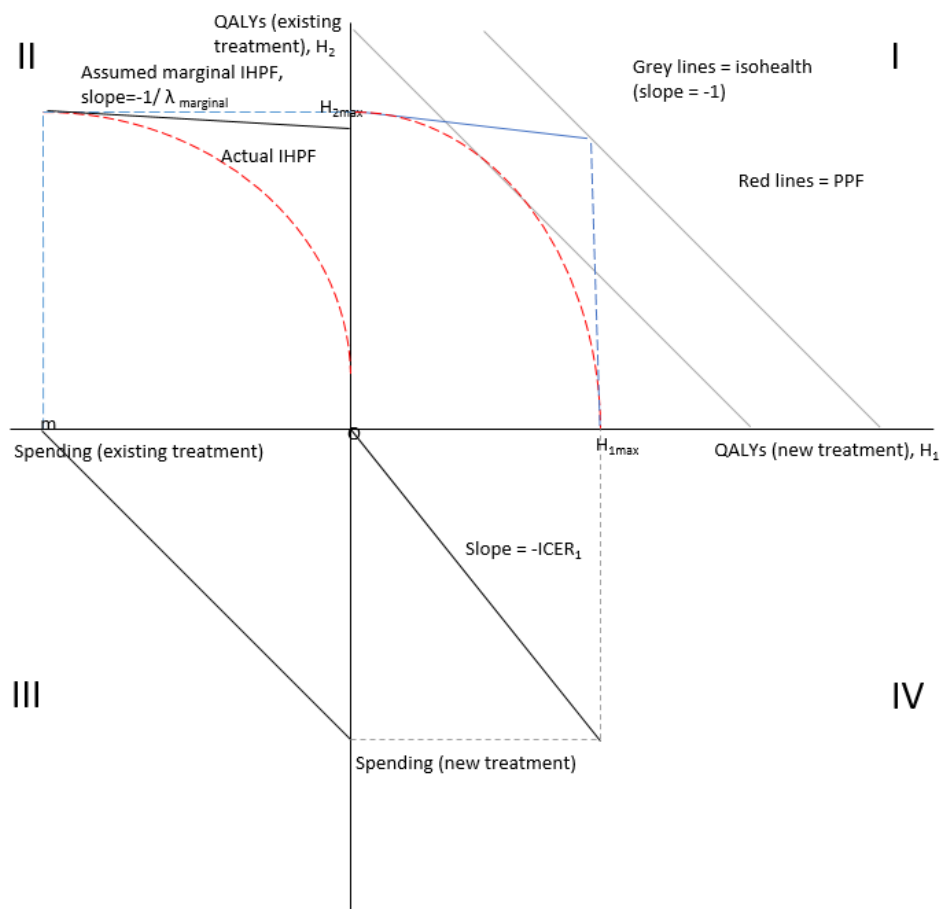


Figure 1: A four-quadrant model of health production

Quadrant I (top-right) presents the final objective function and constraint: maximising total health ( $H_1 + H_2$ ) subject to technological and budgetary constraints. This implies selecting the highest possible level of health (the highest possible isohealth curve, representing  $H_1 + H_2$  and with slope -1) attainable from given current production technologies, prices, and a fixed budget,  $b$  (as illustrated by the given PPF). This quadrant is derived from the three remaining quadrants.

Quadrant IV (bottom-right) displays how much health can be created by given amounts of spending on the new treatment, and has a line with a slope of the negative of the ICER.

Quadrant III (bottom-left) represents the relevant section of the healthcare provider's budget constraint that could potentially be used under conditions of full implementation of the new technology ( $m$ , assumed to be a large proportion of  $b$ )<sup>5</sup>.

Quadrant II (top-left) represents the relevant section of two versions of the IHPF, representing health that is currently generated from spending  $m$  on the most cost-ineffective existing treatment<sup>6</sup>. The non-marginal IHPF for this range of spending (dotted red line) exhibits diminishing returns, reflecting the changes in the cost-effectiveness of health produced by the system when moving away from the margin<sup>7</sup>. The pure marginal IHPF (solid black line), however, is linear, with a slope equal to that of the non-marginal IHPF at its local maximum here ( $-1/\lambda$ ), implying an assumption that healthcare is displaced at a constant opportunity cost equal to the threshold at the margin. Spending the full potential budgetary impact  $m$  on the existing treatment implies a QALY production for this portion of the IHPF of  $H_{2max}$ , consistent with position  $A$  in Quadrant I. The health assumed to be generated and lost at all other points on this IHPF differs according to whether the pure marginal or non-marginal IHPF is used: specifically, the pure marginal IHPF implies a lower opportunity cost in terms of QALYs from existing treatment when additional spending is made on the new treatment, for all levels of budgetary impact strictly greater than 0 and less than or equal to  $m$ .

<sup>5</sup> Note that this does not represent the full budget constraint of the health service, and a full acceptance of this technology (requiring an expenditure of  $m$ ) still leaves  $b - m$  available for spending elsewhere.

<sup>6</sup> Our assumption that displacement of existing healthcare provided occurs in order of least health displaced (i.e., from least to most cost-effective) can be replaced by an assumption that displaced health is displaced in increasing cost-effectiveness, even if this does not represent displacing strictly the least cost-effective treatment first – as is reported empirically by Lomas et al. (2017). This would be one version of a second-best type threshold as characterised by Culyer (2016). In such a situation, the relationship plotted in this quadrant would cease to represent an indirect health production function, and instead represent what might be termed a health displacement function. All further conclusions derived in this paper remain intact, but should be reframed in this context.

<sup>7</sup> In practice, the relevant portion of the IHPF is likely to exhibit a much smaller change in gradient for such a range of spending than that illustrated here. Such an exaggerated change in the slope is displayed for expository purposes.

### 3. Allocation problem

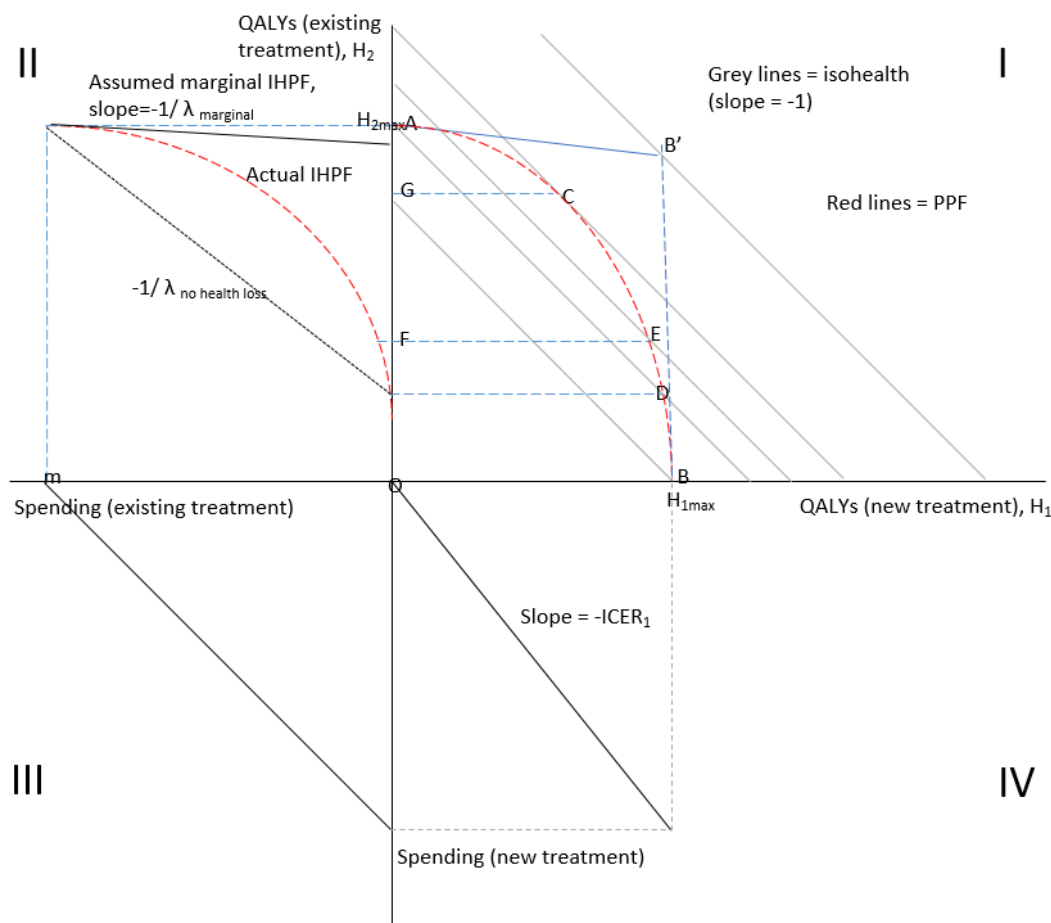


Figure 2: An allocation problem

Point *A* in Figure 2 represents the initial position: no spending can be made on the new, potentially high budgetary impact, treatment before its introduction. The implications of a decision to use the threshold implied by the pure marginal or non-marginal IHPF are illustrated. Assume first that a decision is to be made between accepting the treatment and spending the full amount  $m$  on its provision, implying no spending on the original treatment, or remaining at *A*, with no change in spending – a situation equivalent to a choice involving a potential treatment programme with perfect indivisibility. Under the assumption of pure marginality, acceptance would entail a move from point *A* to point *B'*, consistent with a higher isohealth curve and an associated higher level of overall health, and the acceptance of the technology on these grounds, due to an apparent increase in total health generated. Under the more realistic assumption of non-marginality, however, this entails a move from point *A* to point *B*, consistent with a lower isohealth curve, an associated lower level of overall health, and the rejection of this technology on these grounds, due to a decrease in total health generated.

More interesting cases arise when the treatment is perfectly divisible, and decisions can be made to accept a proportion of spending. Consider the decision to reject this treatment, or accept its full or partial provision. At any point to the right of *A* and the left of *D* (where the PPF cuts the original isohealth line), population health is increased by the partial provision of the new treatment. A discrete choice between rejection and acceptance up to a level just to the left of *D*, for instance *E*, would cause the decision-maker to accept the new treatment and displace  $AF$  QALYs, gaining a greater  $FE$  QALYs, and thus increasing population health by  $FE - AF$ . While this would represent an increase in overall health, it also represents a suboptimal outcome when the objective is the maximisation of total health alone.

Consider point  $C$ . As we move right from point  $A$  along the PPF towards point  $C$ , the slope of the PPF is shallower than that of the isohealth line ( $-1$ ), representing the relative cost-effectiveness of the new treatment compared to existing treatments, and (diminishing marginal) increases in population health from additional spending on the new treatment. As we move right beyond point  $C$ , the converse is true: the slope is steeper than  $-1$ , existing treatment is relatively more cost-effective, and total health is reduced by further use of the new treatment and concomitant displacement of health produced by existing treatments. An optimum in terms of total health is therefore achieved at  $C$ , where the ratio of the marginal health gains of the old and new treatments are equal to 1, the (negative of the) slope of the isohealth line. This implies that, for decisions involving non-marginal impacts on the healthcare system's budget, decision-makers maximising total health should consider the partial acceptance of a new health technology, up to the point that increasing spending on its provision ceases to produce more health than it displaces, rather than necessarily entirely accepting or entirely rejecting the treatment based on its total budgetary impact.

## 4. Implications for horizontal equity

Assuming a constant ICER for the entire population with capacity to benefit, objections to outcome  $C$  may be made on the grounds of a contravention of the principle of horizontal equity – that individuals with equal capacity to benefit should be treated equally, and specifically that individuals with an equal health condition should receive equal healthcare treatment<sup>8</sup>. On such an account, the existence of point  $C$  is irrelevant in a consideration of this treatment in a situation where a greater amount consistent with the selection of point  $B$  could possibly be spent on its provision on equally effective healthcare treatment for the entire patient population<sup>9</sup>. However, this would seem to have substantial echoes of, and form a near-reanimation of, previous related debates (see, *inter alia*, Harris (2005), Claxton and Culyer (2006), Harris (2006), Claxton and Culyer (2007) and Harris (2007)) regarding the principle of decision-making on the basis of cost-effectiveness at all.

Assuming strict exogeneity of the budget constraint, and that the provision of the new treatment does not precisely displace in its entirety the least cost-effective treatment already offered, and no other treatment whatsoever, the provision of this new treatment will displace some, but not the entirety of, one or more programmes of treatment provided elsewhere in the system. In this case, the partial displacement of healthcare elsewhere in the system exacerbates, but does not represent a unique, contravention of the principle of horizontal equity, by denying or delaying treatment to some, but not all, patients in equal need. Indeed, such partial displacement, contravening such a principle of horizontal equity, will occur when any new treatment at all is provided. The acceptance of any new treatment in its entirety, even when its total spending forms only a small proportion of the healthcare budget, will inevitably displace a proportion of spending on healthcare programmes elsewhere in the system.

---

<sup>8</sup> We could, in principle, suggest a conception of horizontal equity that provided each individual with an equal chance of obtaining treatment provided by, for instance, some lottery mechanism which would not be contravened by such an allocation mechanism.

<sup>9</sup> We can illustrate the QALY costs due to such a horizontal equity constraint as  $GC - AG$ , the loss of potential health gains accrued by a failure to partially accept at the optimum.

## 5. Implications for pharmaceutical pricing

The preceding sections have assumed the exogeneity of the price of treatment and of the associated ICER. The decision-making process with regard to low budget impact technologies, where the relevant threshold is known or can be inferred from previous decisions, provides an incentive for the manufacturer to price at the threshold, such that no net health gain is made from the acceptance of the treatment but all surplus is captured by the manufacturer (Gafni and Birch 2006). This nevertheless represents a situation in which the decision-maker, with an objective function of maximising health, and the manufacturer would be unable to effect a mutually beneficial movement from this point.

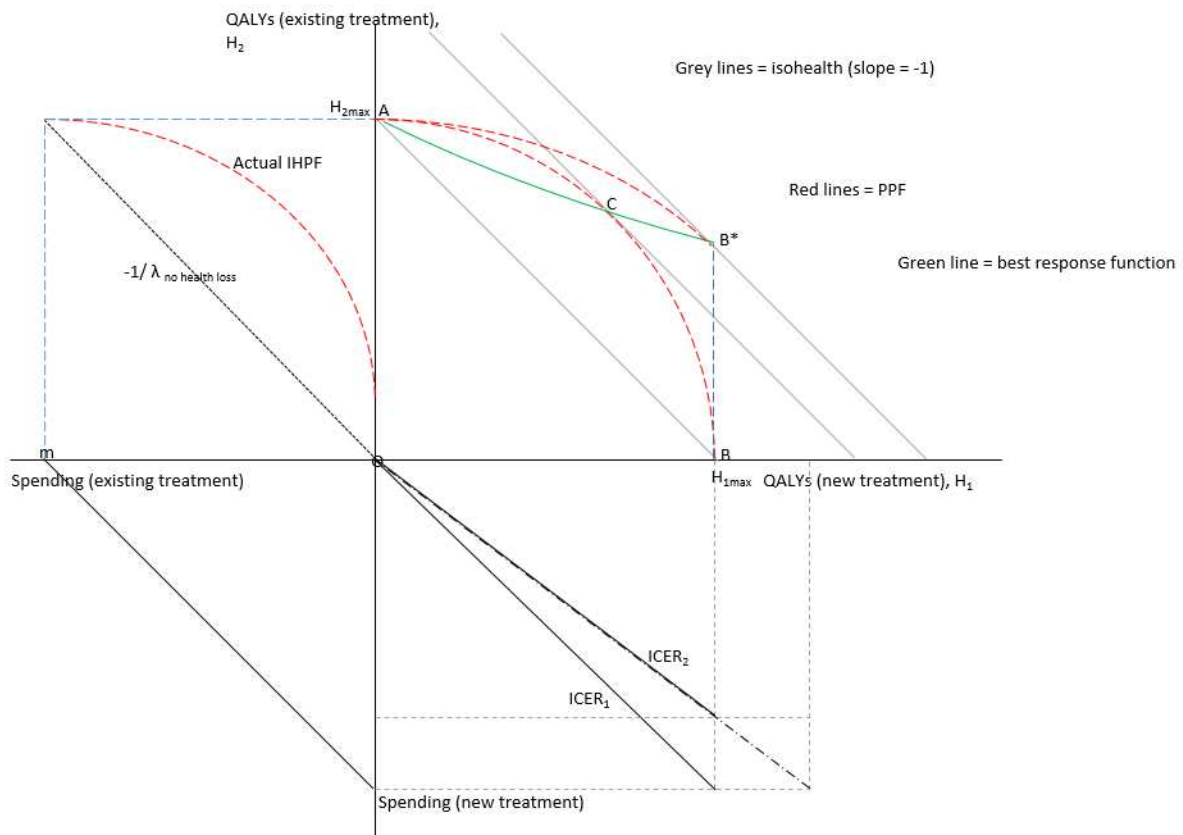


Figure 3: A best response function of the decision-maker

The story is somewhat different in the case of decision-making with large budgetary impacts, where estimates of the shape of the IHPF beyond the margin can be made. Consider now Figure 3, where full implementation does not represent an optimal outcome for a decision-maker trying to maximise total health, but results in zero net health gain (with both A and B lying on the same isohealth curve)<sup>10</sup>. As discussed, holding the ICER as exogenous and given, the requirement to maximise total health would necessitate (setting aside equity concerns) the approval of only a proportion of the total possible volume of use, treating some patients and not others but maximising population health, again at point C. While this partial approval would represent an optimal solution from the point of view of a decision-maker in a one-shot game where price is treated as exogenous, it also represents a point from which there may exist the potential for mutually beneficial gains for both decision-maker and manufacturer. Given the often low marginal cost associated with pharmaceutical products, profit at the margin at this partial

<sup>10</sup> If the manufacturer expects the relevant threshold for a given level of budgetary impact to be set at that involving no total health loss, a profit-maximising manufacturer would, in a one-shot game, set price at this level. We therefore adopt this as a new starting point in this section.

approval point may be positive at a price that the decision-maker would be willing to pay for more units of the treatment. The existence of point  $C$  in such a case would represent a credible threat to manufacturers: the decision-maker can improve population health by not treating all patients with capacity to benefit, and thus reducing manufacturer profits compared to those that can be gained, compared to full implementation with no health loss ( $B$ ). The non-optimality (in the sense that mutual gains between the manufacturer and decision-maker may be possible) of point  $C$  means that a lower price of the treatment (but higher than at the partial implementation point  $C$ ) can result than that which would prevail at point  $B$  where the programme is fully implemented and no net health gains are made. The process by which this may occur is illustrated in Figure 3.

By drawing the implied PPFs resulting from a change in the price (change in the ICER) of a treatment, we can consider one possible price setting outcome. A lower price shifts out the budget constraint from  $ICER_1$  to  $ICER_2$ , leading to a related shift in the PPF. Given that the entire patient population has a total potential benefit of  $H_{1max}$ , the PPF is undefined to the right of this position. At some point, the ICER will fall to a low enough level such that the PPF is tangent to the isohealth curve at a position  $H_{1max}$ : i.e. at point  $B^*$ . Between  $B^*$  and  $A$ , it is possible to plot out a best response function (green line, top-right quadrant) of the decision-maker to a changing ICER, giving a locus of tangencies of the PPF with the isohealth line (including  $C$ ) as this ICER varies. The manufacturer's problem now becomes choosing a price concomitant with a position on this best response curve such that profit is maximised. At  $B^*$  health is maximised at a level at which mutual gains for manufacturer and decision-maker are impossible, with a higher total level of health than at  $C$ , the entire potential patient population treated, and as a result with no horizontal equity concerns arising.

This points to the potential irrelevance of any horizontal equity constraint when the dynamics of price setting are considered. Point  $C$ , held to be in contravention of a particular conception of horizontal equity, does not form an optimal solution but nor does it necessarily form the ultimate solution, but rather it acts as a credible threat that the decision-maker can use when the price for treatment is to be set. If the existence of point  $C$  cannot even be considered by the decision-making body, it cannot exist as a credible threat that may be able to force a move to a lower price (lower ICER) for the treatment that can also result in the full population receiving treatment, with no potential horizontal equity concerns arising in this final outcome. While manufacturer knowledge of the decision-making process for health technology assessment with a constant threshold implied by the IHPF at the margin leads to a situation where net health gains are driven towards zero and all surplus captured by the manufacturer, knowledge of the existence of the non-linear PPFs (and the associated best-response functions of the decision-maker) may lead to a situation where the price set by the manufacturer is that associated with the maximisation of total health, with a positive net health gain over non-implementation.



## 6. Conclusion

The conventional approach of rejecting or accepting a new healthcare technology by comparison of the ICER to some cost-effectiveness threshold is appropriate, as an approximation to a full mathematical programming solution to the maximisation of total health, if the impact on the healthcare service's budget is marginal. Such a case generates no concerns regarding horizontal equity, as new treatments are either accepted or rejected in their entirety. This paper presents a simple theoretical model illustrating the choice faced by decision-makers in situations where the budgetary impact of an intervention is great enough to make the non-linearity of the relevant section of the IHPF important, and where those decision-makers have knowledge of the shape of the IHPF as we move away from the margin. In such a situation, optimal acceptance of the new intervention in a static exogenous-ICER case may be partial rather than full. If implementation is partial rather than full, horizontal equity concerns may (appear to) arise from the fact that patients with identical health conditions and equal capacity to benefit will not receive equal treatment. It is important, however, to recognise that such horizontal equity issues exist within the actually existing system, and are exacerbated when any new treatment is accepted and health provided by other existing treatments in other sectors is displaced. Furthermore, even if the relevance of this horizontal equity constraint is accepted in the statics, it does not exist in a likely price setting game that arises from treating the ICER as endogenous rather than exogenous. Such a price setting process may result in an ICER that is lower than that implied by a threshold involving no loss of health ( $\lambda_{\text{no health loss}}$ ), and which may maximise population health while treating the entire population with capacity to benefit.

## References

- Claxton, K. and A. J. Culyer (2006). 'Wickedness or folly? The ethics of NICE's decisions'. en. *Journal of Medical Ethics*, 32: 373–377. ISSN: 0306-6800, 1473-4257. DOI: 10.1136/jme.2006.016204. URL: <http://jme.bmj.com/content/32/7/373> (visited on 22/02/2017).
- Claxton, K. and A. J. Culyer (2007). 'Rights, responsibilities and NICE: a rejoinder to Harris'. en. *Journal of Medical Ethics*, 33: 462–464. ISSN: 0306-6800, 1473-4257. DOI: 10.1136/jme.2006.018903. URL: <http://jme.bmj.com/content/33/8/462> (visited on 22/02/2017).
- Claxton, K., S. Martin, M. Soares, N. Rice, E. Spackman, S. Hinde, N. Devlin, P. C. Smith and M. Sculpher (2015). 'Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold'. eng. *Health Technology Assessment (Winchester, England)*, 19: 1–503, v–vi. ISSN: 2046-4924. DOI: 10.3310/hta19140.
- Culyer, A. J. (2016). 'Cost-effectiveness thresholds in health care: a bookshelf guide to their meaning and use'. *Health Economics, Policy and Law*, 11: 415–432. ISSN: 1744-1331, 1744-134X. DOI: 10.1017/S1744133116000049. URL: <https://www.cambridge.org/core/journals/health-economics-policy-and-law/article/cost-effectiveness-thresholds-in-health-care-a-bookshelf-guide-to-their-meaning-and-use/B7DE29D8D030FEA9B122869088AC3E14> (visited on 25/10/2017).
- Drummond, M. F., M. J. Sculpher, K. Claxton, G. L. Stoddart and G. W. Torrance (2015). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford University Press. (Visited on 23/02/2017).
- Epstein, D. M., Z. Chalabi, K. Claxton and M. Sculpher (2007). 'Efficiency, Equity, and Budgetary Policies: Informing Decisions Using Mathematical Programming'. en. *Medical Decision Making*, 27: 128–137. ISSN: 0272-989X. DOI: 10.1177/0272989X06297396. URL: <http://journals.sagepub.com/doi/abs/10.1177/0272989X06297396> (visited on 23/02/2017).
- Gafni, A. and S. Birch (2006). 'Incremental cost-effectiveness ratios (ICERs): The silence of the lambda'. *Social Science & Medicine*, 62: 2091–2100. ISSN: 0277-9536. DOI: 10.1016/j.socscimed.2005.10.023. URL: <http://www.sciencedirect.com/science/article/pii/S0277953605005496>.
- Harris, A. H. (2016). 'Beyond the threshold'. eng. *Health Economics, Policy, and Law*, 11: 433–438. ISSN: 1744-134X. DOI: 10.1017/S1744133116000050.
- Harris, J. (2006). 'NICE is not cost effective'. en. *Journal of Medical Ethics*, 32: 378–380. ISSN: 0306-6800, 1473-4257. DOI: 10.1136/jme.2006.016691. URL: <http://jme.bmj.com/content/32/7/378> (visited on 22/02/2017).
- Harris, J. (2005). 'It's not NICE to discriminate'. en. *Journal of Medical Ethics*, 31: 373–375. ISSN: 0306-6800, 1473-4257. DOI: 10.1136/jme.2005.012906. URL: <http://jme.bmj.com/content/31/7/373> (visited on 22/02/2017).
- (2007). 'NICE rejoinder'. en. *Journal of Medical Ethics*, 33: 467–467. ISSN: 0306-6800, 1473-4257. DOI: 10.1136/jme.2007.020628. URL: <http://jme.bmj.com/content/33/8/467> (visited on 22/02/2017).
- Lomas, J., K. Claxton, S. Martin and M. Soares (2017). 'Resolving the 'cost-effective but unaffordable' paradox': estimating the health opportunity costs of non-marginal changes in available expenditure'. *[under review, Value in Health]*.
- McCabe, C., K. Claxton and A. J. Culyer (2008). 'The NICE cost-effectiveness threshold: what it is and what that means'. eng. *Pharmacoeconomics*, 26: 733–744. ISSN: 1170-7690.
- National Institute for Health and Care Excellence (2015a). *Daclatasvir for treating chronic hepatitis C*. eng. URL: <https://www.nice.org.uk/guidance/ta364> (visited on 16/05/2017).
- (2015b). *Ledipasvir-sofosbuvir for treating chronic hepatitis C*. eng. URL: <https://www.nice.org.uk/guidance/ta363> (visited on 16/05/2017).
- (2015c). *Ombitasvir-paritaprevir-ritonavir with or without dasabuvir for treating chronic hepatitis C*. eng. URL: <https://www.nice.org.uk/guidance/ta365> (visited on 16/05/2017).
- Paulden, M. (2016). 'Opportunity cost and social values in health care resource allocation'. English. PhD thesis. University of Alberta. URL: <https://doi.org/10.7939/R3M902D4P>.
- Paulden, M., J. O'Mahony and C. McCabe (2017). 'Determinants of Change in the Cost-effectiveness Threshold'. eng. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, 37: 264–276. ISSN: 1552-681X. DOI: 10.1177/0272989X16662242.
- Stinnett, A. A. and J. Mullahy (1998). 'Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis'. eng. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, 18: S68–80. ISSN: 0272-989X. DOI: 10.1177/0272989X98018002S09.