Theoretical models of the cost-effectiveness threshold, value assessment, and health care system sustainability

March 2018

INSTITUTE OF HEALTH ECONOMICS
ALBERTA CANADA

UNIVERSITY OF ALBERTA

INSTITUTE OF HEALTH ECONOMICS

The Institute of Health Economics (IHE) is an independent, not-for-profit organization that performs research in health economics and synthesizes evidence in health technology assessment to assist health policy making and best medical practices.

IHE BOARD OF DIRECTORS

Chair

Dr. Lorne Tyrrell – Professor & Director, Li Ka Shing Institute of Virology, University of Alberta

Government and Public Authorities

Mr. Justin Riemer – Assistant Deputy Minister, Alberta Health

Mr. Jason Krips – Deputy Minister, Economic Development and Trade

Mr. Tim Murphy – VP Health, Alberta Innovates

Dr. Kathryn Todd – VP Research, Innovation & Analytics, Alberta Health Services

Academia

Dr. Walter Dixon – Interim VP Research, University of Alberta

Dr. Jon Meddings – Dean of Medicine, University of Calgary

Dr. Richard Fedorak – Dean of Medicine & Dentistry, University of Alberta

Dr. Ed McCauley – VP Research, University of Calgary

Dr. Neal Davies – Dean of Pharmacy & Pharmaceutical Sciences, University of Alberta

Dr. Braden Manns – Svare Chair in Health Economics and Professor, Departments of Medicine and Community Health Sciences, University of Calgary

Dr. Rick Szostak – Chair, Department of Economics, University of Alberta

Industry

Mr. Robert Godin –Director, Market Access Strategy & External Relations, AstraZeneca

Ms. Jennifer Chan – VP, Policy & Communications, Merck Canada

Ms. Franca Mancino – Head of Market Access, Sanofi Canada

IHE

Mr. Doug Gilpin – Chair, Audit & Finance Committee

Dr. Christopher McCabe – Executive Director & CEO, Institute of Health Economics

Ms. Allison Hagen – Director of Finance, Operations & Administration, Institute of Health Economics

IHE White Paper

Theoretical models of the cost-effectiveness threshold, value assessment, and health care system sustainability

Prepared by:

Himani Pandey, Department of Economics, University of Alberta

Mike Paulden, School of Public Health, University of Alberta

Christopher McCabe, Institute of Health Economics

Acknowledgements

The authors are grateful to:

- Dr Karl Claxton and Mr. James Lomas for helpful discussions on the material presented in this report. The views expressed in this report are the authors alone.

Corresponding Author

Please direct any inquiries about this report to Christopher McCabe, cmccabe@ihe.ca.

Funding

This report was supported by a financial contribution from Patented Medicines Price Review Board. The completed report was submitted to Patented Medicines Price Review Board in March 2018.

Declared Competing Interest of Authors

Competing interest is considered to be financial interest or non-financial interest, either direct or indirect, that would affect the research contained in this report or create a situation in which a person's judgement could be unduly influenced by a secondary interest, such as personal advancement.

The authors of this publication claim no competing interest.

Suggested Citation

Pandey H, Paulden M, McCabe C. *Theoretical models of the cost-effectiveness threshold, value assessment, and health care system sustainability*. Edmonton (AB): Institute of Health Economics; 2018.

Reproduction, redistribution, or modification of the information for any purposes is prohibited without the express written permission of the Institute of Health Economics

Institute of Health Economics, 2018 www.ihe.ca

Abbreviations

All abbreviations that have been used in this report are listed here unless the abbreviation is well known, has been used only once, or has been used only in tables or appendices, in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.

CEA	cost-effectiveness analysis
DALY	disability-adjusted life year
GDP	gross domestic product
ICER	incremental cost-effectiveness ratio
k	supply-side threshold
λ	threshold
PCT	primary care trust
PMPRB	Patented Medicine Price Review Board
QALY	quality-adjusted life year
R&D	research and development
v	demand-side threshold
VSL	value of a statistical life
WTP	willingness-to-pay

Table of Contents

Acknowledgements i

Abbreviations ii

SECTION 1: Supply- and demand-side models of the cost-effectiveness threshold; budget impact and financial sustainability 5

Changing value of threshold 6

Budget impact 6

Summary 7

SECTION 2: The relationship between demand- and supply-side cost-effectiveness thresholds 8

Linking the two sides conceptually 8

Link between v and k 9

Figure 1: The relationship between v and k cost-effectiveness thresholds 10

Conceptual background 10

Alternative scenarios 11

Factors affecting the threshold 12

Comparative statics 12

Comparative dynamics 13

Summary 14

SECTION 3: Empirical estimates of demand- and supply-side cost-effectiveness thresholds 15

Table 1: Summary of studies estimating the relationship between health expenditures and health outcomes .... ***Error! Bookmark not defined.***

Summary 17

SECTION 4: Incorporating equity weights into consideration of value-based pricing: evidence and issues 21

Summary 22

SECTION 5: Exposition of mechanisms for incorporating societal contribution to global pharmaceutical research and development into value-based pricing assessments 23

Figure 2: Consumer and producer threshold curves, reflecting the relationship between the threshold (λ), net population benefit (consumer surplus), and manufacturer profit (producer surplus) 24

Table 2: Optimal threshold (λ) or range containing optimal threshold, for each objective* 26

SECTION 6: Summary 27

Policy considerations 28

References 30

SECTION 1: Supply and demand-side models of the cost-effectiveness threshold, budget impact, and financial sustainability

Cost-effectiveness analysis (CEA) aims to assess the value of health gain attributable to a technology in order to assist the decision-makers in allocating scarce resources efficiently.[1] These analyses express costs in monetary units (such as dollars) and health gains in units of health such as quality-adjusted life years (QALYs) gained or disability-adjusted life years (DALYs) averted. Incremental cost-effectiveness ratios (ICERs) are calculated, which reflect the health gains compared to the costs incurred in providing the healthcare intervention. However, ICERs alone are not sufficient to inform policy decisions. A reference value is required to establish whether the ICER for a specific technology represents good or poor value;[2] this value is often referred to as the *cost-effectiveness threshold*. The threshold symbolizes a cut-off point or a *critical ratio*, as defined by Weinstein and Zeckhauser,[3] for allocating resources among competing uses in a budget-constrained environment.

There are multiple ways to establish the threshold value. There is substantial disagreement among authors on the most appropriate means for identifying the threshold and what value it should take.[4] Garber and Phelps[5] describe the threshold as essentially a value judgment that depends upon several factors, including who the decision-maker is, what the purpose of the analysis is, how health gains and costs are valued (and weighted), what risks are considered, and, finally, what resources are available. The answers to these factors essentially indicate the perspective considered while formulating the threshold; that is, which stakeholders' (patient/consumer, insurer, government) value for health and cost is reflected by the threshold. A consumer's threshold should reflect their willingness to pay, an insurer's threshold would reflect the market demand for the intervention, and a government's threshold should reflect the social value consensus, perhaps through a social welfare function.

The diverse ways in which a threshold can be estimated can affect how a decision-maker interprets them and hence uses them for decision-making. According to McCabe et al.,[6] a threshold can be defined in three potential ways: i) it could be inferred from previous decisions; ii) it could be defined to set an optimal healthcare budget; and iii) it could be set to exhaust an exogenously given budget.

The first method involves using benchmark interventions or previous decision rules to guide current decisions. This was originally proposed by Weinstein and Zeckhauser.[3] These thresholds emerge from retrospective analysis of an existing practice.[7] An example could be the threshold of US$50,000 per QALY gained, previously used in the United States (now increased to US$100,000 per QALY), which is believed to be based on the CEA of the dialysis for chronic renal failure.[8] The use of inferred thresholds may be convenient but, due to issues of unknown confounders, they may be too high or too low; that is, they may not adequately reflect either society's willingness-to-pay (WTP) or the opportunity costs of adopting the technology.

The second and third methods are popularly cited as the *demand-side* and *supply-side* approaches, respectively. The demand-side approach requires the society's WTP for health care to determine the threshold which would guide the healthcare budget accordingly. The marginal WTP can be elicited through several ways: i) using revealed/stated preference methods through a representative sample; ii) contingent valuation studies by using value of health/life employed in other areas of resource allocation; and iii) assuming the gross domestic product (GDP) per capita would reflect it.[6, 9] The latter approach has been heavily criticized for the implicit assumption that there is a fixed relationship between GDP and the appropriate magnitude of expenditure on health care, whereas this is a policy decision that can legitimately vary in terms of whether there is a relationship and, if so, the form and magnitude of that relationship.

The WTP in general is hard to quantify and generalize, as individuals may attach different weights to benefits from health care depending upon the value they attach to health, the process of health care, the other economic activities that health is an input to, and their attitude to risk. Hence, forming a social WTP by aggregation can be both empirically complex and conceptually difficult to justify. In general, demand-side methods fail to reflect affordability of new technologies, and hence do not inform the real trade-offs that are at the center of health technology reimbursement decision-making.[8]

The supply-side method is conceptually related to the *cost-effectiveness league table* approach, wherein the interventions are ranked in increasing order of their ICERs, and funding decisions are made starting from the one with the lowest ICER moving upward until the budget is exhausted.[2] However, the league table approach does not necessarily throw light on other issues such as equity (size and characteristics of the affected population), ethical concerns (provision of life-saving drugs and treatments), and political feasibility.[6] It also assumes that that all interventions are ranked in correct order, are divisible, exhibit constant returns to scale, and are completely independent of each other. In the event of uncertainty, this method fails to deliver a first-best solution.[10] The empirical work of Claxton et al.[11] shows that, when we only assume a fixed budget, the displaced technologies are not necessarily the least productive ones; rather, they may be those that are managerially the most convenient in the short run to remove or reduce. In such a case, the threshold should reflect the ICER of the displaced technology, which will be a second-best solution and higher than the first-best.

Changing value of threshold

In the supply-side model, it is important to note that the threshold is not constant but rather will change in response to a number of factors. Paulden et al.[12] present comparative static analyses that show how changes in budgets, demand for currently funded technologies, medical inflation, and pace of innovation will lead to changes in the supply-side cost-effectiveness threshold. The threshold can increase (decrease) with budget expansion (contraction), increasing (decreasing) demand of existing technologies and decreasing (increasing) effectiveness of existing technologies. The dynamic nature of the cost-effectiveness threshold needs to be born in mind when considering its use in a value-based pricing framework. Fixing the threshold value for a period of time requires the capacity compensating changes in budget to maintain the threshold or the actual impact of a value-based price may diverge from the intended objective.

Budget impact

In determining a supply-side threshold, a candidate technology can have an impact on the budget that may result in displacement of one technology (marginal budget impact) or multiple technologies (large/non-marginal budget impact). This can have repercussions in empirical estimation of a supply-side threshold. The more technologies that get displaced, the lower the ICER threshold is going to be, such that the threshold will reflect a weighted average of the displaced technologies.[12]

A detailed discussion on effectiveness versus affordability is provided by Lomas et al.,[13] who have examined the opportunity cost of larger scale (non-marginal) budget impacts. They incorporated the non-marginal budget impacts while estimating the supply-side threshold and showed that, as theory predicted, the threshold is indeed lower than that found by Claxton et al.[11] This reiterates the observation from Paulden et al. that a technology may be affordable (as in its budget impact is less than the total budget from which it will be funded), but not cost effective.[12]

Summary

CEAs produce ICERs and an input to resource allocation decision processes. However, an ICER cannot be interpreted without a reference standard that differentiates good from poor value. This reference standard is called the *cost-effectiveness threshold*, and it is an empirical parameter for decision-makers that may be estimated either as the WTP for an additional unit of health benefit, when the healthcare budget is assumed to be unconstrained, or as the opportunity cost of adopting a technology at the margin of the healthcare budget, when the budget is assumed to be constrained. As healthcare budgets are increasingly recognized as being at least partially constrained, the latter definition – referred to as the *supply-side* model of the cost-effectiveness threshold – is broadly considered the appropriate form of the cost-effectiveness threshold to use in decision-making. In the context of the Patented Medicine Price Review Board (PMPRB), the supply-side cost-effectiveness threshold is an operationalization of the concept of an excessive price that defines a price as "excessive" if it displaces more benefit through opportunity cost than the technology produces.

The supply-side cost-effectiveness threshold can be estimated empirically, as has been demonstrated by Claxton and colleagues for a range of countries, using both QALYs and DALYs as the measure of value. The dynamic characteristics of the supply-side cost-effectiveness threshold are increasingly well understood, which allows the policy implications of any specific value-based pricing policy to be well understood. Given the advances in the theoretical and empirical evidence on the supply-side cost-effectiveness threshold, its use has the potential to add to the transparency of PMPRB processes, and is consistent with an evidence-based policy paradigm.

SECTION 2: The relationship between demand- and supply-side cost-effectiveness thresholds

The cost-effectiveness threshold or the cut-off point used to make critical funding decisions in a budget-constrained environment can be conceptually viewed through two perspectives, as discussed in Section 1. The demand-side perspective encapsulates the society's WTP for health gains or avoiding health losses. The supply-side perspective, on the other hand, captures the opportunity cost that results from disinvestment to fund more cost-effective technologies.

The demand-side threshold (v) is determined by the willingness to contribute to the private health expenditure by individuals in a society. The WTP of every individual can be aggregated to ascertain this value or, in some cases, the collective WTP for the society's health can be evaluated as well. This value is therefore linked to the incomes and share of income devoted to health by individuals.

The supply-side threshold (k) is determined by the opportunity costs of disinvesting when new interventions need funding. Since this evaluation is done in a budget-constrained environment, the healthcare budget has a pivotal role in affecting this threshold value. The healthcare budget is a decision undertaken by the government in power, which should ideally reflect the preferences of the people that elected it. In that way, the individual's income again affects the threshold value through the contribution to tax revenue, which the government uses for funding various sectors of the economy.

Linking the two sides conceptually

The demand-side considers the threshold as the WTP for health improvement by individuals or the consumption value of health. Using this definition, health becomes a consumption good for individuals, entering their utility function. The objective is, therefore, to maximize the utility over health and other consumer goods given the budget of individuals. This maximization exercise generates a demand for health as a consumption good if prices for health and other goods are given and income is exogenously determined. This "demand for health" is the threshold from the demand-side.

Factors affecting the demand-side threshold are:

- individual income;
- prices of health services;
- prices of substitutes and complementary goods to health; and
- environmental changes affecting the demand exogenously.

The supply-side considers the opportunity cost of the government budget spent on health, that is, the marginal productivity of the health care system, which can be visualized as an input in the health production function of the government.

The government or the public sector is believed to contribute to the total welfare of the individuals using investments in health care, education, infrastructure, et cetera as inputs. Health is therefore considered as an investment good from this side of the economy that can lead to higher income in the economy. Tax revenues are the income for the government in this model. The objective function for the government is assumed to be to maximize the net benefits (profit function) given the input prices (for example, costs of health, education), or to minimize the costs of producing the government's total contribution to welfare given the tax revenue.

The outcome would be the optimum amount of health care that the government would generate/supply such that total welfare of individuals is maximized. This amount would reflect the marginal productivity of health care and will account for the competing uses of the government budget. Therefore, it reflects the supply-side cost-effectiveness threshold.

Factors affecting the supply-side threshold are:

- prices of healthcare technologies;
- prices/cost of providing other services by government;
- tax revenue collected by government; and
- technological developments in health sector.

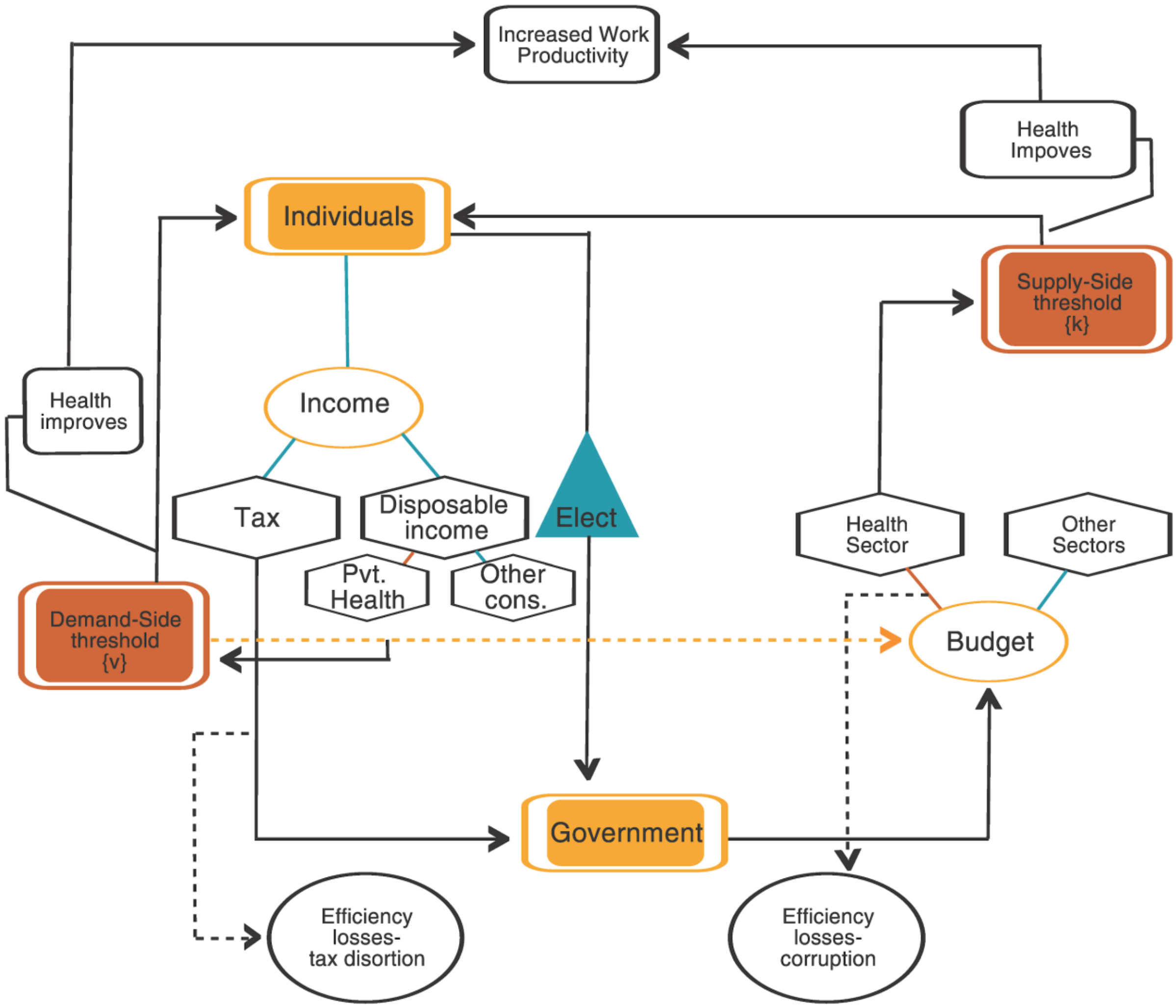
Factors affecting the demand-side and supply-side thresholds, as discussed above, are interrelated or common. Thus, both the measures can change simultaneously. The income base for individuals and the government play an important role in determining how the two measures are related.

Link between *v* and *k*

It is often argued that for making funding decisions, the supply-side threshold (*k*) is what should be considered because it reflects opportunity costs. An estimate of the demand-side threshold (*v*), which reflects the consumption value of health or how much individuals are willing to pay for mortality reductions, *can* reflect the social value of health and thus inform decisions about scale of resources that should be allocated to the healthcare budget; *v* may therefore reflect the size of the healthcare budget.[10]

However, this link is dependent upon the decisions of government in fixing the healthcare budget. The government's capacity and willingness to reflect society's preferences for health in fixing its budget is a dimension of the political economy models. Further, government's scale of activities can impact the efficiency in health delivery, and therefore understanding the efficiency losses and gains of government activity is critical to determine the relationship between *k* and *v*. A conceptual model is outlined in Figure 1.

FIGURE 1: The relationship between *v* and *k* cost-effectiveness thresholds



Conceptual background

Through a simple flow mechanism, the basic working of this threshold framework can be discussed. Under varied assumptions, the tax rate, share of contribution to private health (WTP for better health by individuals), and share of contribution to public health (government healthcare budget) are the main factors that can affect the threshold value.

The demand-side threshold (*v*) is determined through the amount individuals are willing to spend on private health out of their disposable income (income after paying taxes). The supply-side threshold (*k*) is determined through the amount government allocates towards health care out of the total budget (based on total tax revenues collected). Taxation appears to be an important factor on both sides of the estimation. Therefore, efficiency losses are important in determining both *v* and *k*. For *v*, there are losses due to distortionary taxes, while, for *k*, welfare losses occur due to misuse and misallocation of the tax revenues.

The role of the electoral system is crucial in this model. Individuals elect a government assuming the government would reflect on their preferences, and thus information flow plays an important role. The closer the government reflects the individual preferences of voters, the smaller the expected divergence between *v* and *k*, all things being equal.

Alternative scenarios

When v *and* k *are equal*

Both measures are believed to be equal only under the prevalence of strict assumptions of a hypothetical world where there is full information for individuals and the government regarding individual preferences, and the government has the capacity to act upon them. This would depend in part upon a transparent electoral system, an efficient health care system, and trust in the public sector. At the same time, all individuals must have homogeneous preferences such that the social-welfare function can truly reflect on the welfare of everyone. Under heterogeneity of preferences, it is likely impossible to aggregate individual preferences in a way considered socially legitimate, and hence any particular allocation of resources to health care (determining *k*) is unlikely to robustly align with the preferences of the majority of individuals (determining *v*).

When v *and* k *differ*

The real world is far from this perfect scenario. There are informational constraints and often a society's preferences for health will not be reflected in the healthcare budget. In such a scenario, a political economy model reflecting the trust individuals place in their governments can throw light on this issue. If the individuals in a society have faith that the taxes paid by them are efficiently used up for providing for their health care and other services; and the level of satisfaction from the public health care system is high, then a demand-side threshold may be similar to the supply-side threshold. However, significant dissatisfaction with the public health care system will cause the demand-side threshold to be higher than the supply-side threshold, as those with the ability to pay purchase supplementary health insurance to meet their expectations.

It is likely that *v* is greater than *k* when the basic premise of trust in government is not strong enough. It can be routed in subtle ways through the welfare losses that occur due to the taxation system, and if the WTP of individuals is greater for themselves compared to that of the society in general. This creates uncertainty in the economy, and efficiency losses will increase as a result. It is worth noting that with efficiency losses in place, both due to deadweight losses in taxation and from the government's performance in allocating the budget, the value of *k* will need to be driven higher than the original *v* in order to compensate and match the social preferences. This is unlikely to be plausible with limited budgets.

Moreover, the value of *v* can be higher than *k* because the government operates with economies of scale. The value of *v* is affected by the coverage of the private insurance companies, and is therefore limited by the people insured in the economy either through employment or otherwise. The value of *k*, on the other hand, is governed by public insurance, which has a much wider coverage, potentially driving down the average administrative costs. The wider coverage can allow the government to provide various health technologies/drugs at a lower costs, due to greater negotiating power and economies of scale in associated clinical services compared to private providers, thus making it more cost-effective. This can have an effect of reducing the value of *k* per se. The extent to which this effect is observed in practice will depend heavily upon the degree to which government can realize the economies of scale and scope.

The other case of *k* being greater than *v* is plausible when the investment value of health is very high, and the individuals have high levels of certainty regarding the intentions of the government. Cuba is an example of this, where the government is highly committed to investment in health, thus making the individuals lower their own WTP for health care.

The ratio of *k* and *v* is also likely to differ across countries or within a country's provinces, owing to different income levels. Healthcare spending is assumed to have positive income elasticity with respect to incomes, and therefore higher income regions are expected to have higher healthcare budgets and greater spending. Also, in lower income countries, the size of the healthcare budget is likely to be constrained by the operational ability of countries to raise tax revenues.[15]

Factors affecting the threshold

Comparative statics

Assumptions

- Individuals have homogeneous preferences.
- Income is exogenously determined.
- An increase in private or public health spending is assumed to have an equivalent effect in improving health.

Given the above assumptions, the threshold change is analyzed under different circumstances holding everything else static or constant.

Increase in tax rate

Taxes play an important role in generating revenue for the government and thus in deciding the budget available for funding health interventions. At the same time, it also determines the income at the disposal of individuals. Therefore, the budget of both individuals and the government is affected through the tax rate. Under a static scenario with all else being equal, a change in tax rate can change both valuations of the threshold (*k* and *v*). As the tax rate rises, and tax revenue for the government rises, then, all else being equal, the public allocation to health increases and the supply-side threshold (*k*) rises. At the same time, with a higher tax rate, and less disposable income with individuals, their WTP for health may fall if the health delivery system is believed to be efficient, and thus the demand-side threshold (*v*) would fall. The impact on total health expenditure remains ambiguous, as the direction of dominance between public and private health expenditure needs to be evaluated.

However, if government as discussed above is believed to have a wider coverage for healthcare provision and the private WTP is only affected by the availability of private insurance, which is limited in reach due to its availability to individuals with employment or higher income or risk averse nature, then a rise in *k* could dominate the fall in *v*, thus causing total health expenditure and health in general to rise.

Increase in share of private expenditure on health

The share of the contribution to health expenditure by individuals from their disposable income is key to determine their WTP for own health, thus deciding the value of *v*. A higher contribution to health expenditure by individuals, either due to greater private insurance or better employment benefits, would cause the value of *v* to rise, all else being equal. This would be independent of the tax revenues, and so there would be no impact on the value of *k*. A higher *v* with a constant *k* would unambiguously cause the total population health to improve, due to the increase in total health expenditure.

Increase in share of public revenues allocated to health care

The government allocation of its budget to healthcare investment is an important determinant of k. All else being equal, with greater public spending on health, the individual WTP for health would either remain same or fall, thus causing k to rise with a lower or unchanged v. Total health expenditure would increase, again leading to improvement in total population health.

Comparative dynamics

The analysis can be expanded to consider a dynamic situation, where the impact of a shock can be felt over multiple periods like a multiplier effect. Considering dynamic effects is important, as health improvements are linked to the earning capacity of individuals through their ability to work more efficiently with better health. Income therefore becomes endogenous to health, and we would expect a multiplier effect to be observed.

Assumptions

- Income is endogenous to health levels.
- Expenditures in health, whether private or public, are positively related (linearly) to health outcomes.
- Individuals have homogeneous preferences.

All dynamic effects will be analyzed over two periods (which can easily be extended to further periods), modelling the effects as a chain reaction.

Increase in tax rate

In the first period, a rising tax rate would cause the government budget to rise and the individual disposable income to fall. Thus, the value of k increases and value of v decreases. This would cause the private health expenditure to fall and the public health expenditure to rise. The impact on total health expenditure is the same as discussed in the *Comparative statics* section above. There is ambiguity in general; however, there are reasons to expect that health expenditure may rise if public expenditure dominates private expenditure, causing the total population health to improve.

Similarly, in the second period, if there is ambiguity then the further chain reaction cannot be ascertained without simplifying assumptions. If the government operates with economies of scale due to wider coverage of health care, then there would be an unambiguous rise in incomes with higher health expenditures. This would lead to higher v and higher taxes as well, which could convert to higher k. This would lead higher levels of health in all periods thereafter.

Increase in share of private expenditure on health

In the first period, higher private expenditure on health would cause v to rise with a constant k. There would be an unambiguous rise in total health expenditure and total population health.

In the second period, this would cause the incomes of individuals to rise with enhanced work productivity, thus leading to rising taxes and contribution to private health. As a result, k and v would rise, causing total health expenditure and health levels in the society to increase in the periods thereafter, continuing the multiplier effects.

Increase in public allocation on health

In the first period, an increase in the government's healthcare budget will, all else being equal, lead to a rise in *k* and a fall/no-change in *v* (depending on the level of substitutability of private and public expenditure in health). Total health expenditure will rise, leading to a rise in health levels in society.

Thus, in the second period, with higher individual incomes, the multiplier effect can set in and the values of *k* and *v* can continue to rise to match social preferences and the budget, leading to higher levels of health.

Summary

Having discussed the conceptual background to the possible links between demand- and supply-side thresholds (*v* and *k*, respectively) with an outlook on how shocks such as changes in tax rate, private provision to health care, or public allocation to health care can impact the overall health levels in the economy in the short run as well as longer time periods, the various channels of threshold determination are understood. This working knowledge can be improved further by considering the political economy and allowing the relaxation of excessively simple assumptions, such as a linear relationship between health expenditures and health outcomes, and homogeneity in individual preferences. In the real world, we witness diminishing returns to health expenditures and thus non-linear relationships, which would reduce the multiplier effects eventually. Further, it is well known that individuals are heterogenous and their preferences can also vary with respect to their incomes and other factors.

In general, the potential for greater efficiency through economies of scale suggest that *k* can be lower than *v*, while preference heterogeneity means we cannot conclude as to whether *v* can be reliably identified and hence whether any specific value of *k* is above or below the *v* that would be observed if all health care was funded through private finance (either out-of-pocket or private insurance). The analysis presented in this section would suggest that, if decisions based on *k* were substantially discrepant with individuals' values, there would be an increasing use of private insurance. As the proportion of Canadian health care that is privately funded has been stable at 30% since the year 2000, with around 12% of care being funded by private insurance, there is a prima facie case that *k* is sufficiently close to *v* for the majority of Canadians.

SECTION 3: Empirical estimates of demand- and supply-side cost-effectiveness thresholds

As discussed elsewhere, a cost-effectiveness threshold can be conceptually understood through two approaches: i) how the society values health gains, or/and ii) the opportunity cost involved in displacement of another technology in the health care system.[16] The literature on the estimation of cost-effectiveness thresholds are therefore categorized widely into these two domains, also referred to as the *demand-side* and *supply-side*, respectively.

In order to conduct a critical review of the related literature, a systematic search was conducted using the key papers in this field as identified by Claxton et al.,[11] dating up until 2013. We performed a forward and backward search on those key initial papers, updating up until 2017, to find 67 related papers, including some very recent systematic reviews.[*] A recent review by Vallejo-Torres et al.[17] identified 38 studies, out of which 29 were classified as demand-side studies and the remaining nine as supply-side studies. Their review findings suggest that estimates based on the demand-side approach tend to be higher than the estimates based on the supply-side approach. This, according to the authors, could suggest that "some interventions with positive social net benefits, as informed by individuals' preferences, might not be an appropriate use of resources under current budget constraints."

The literature on threshold estimation through the demand-side includes a wide variety of studies that allow for eliciting social preferences on the threshold via an aggregated method of adding the ratios of WTP over QALY gains, or via a disaggregated method of computing a ratio from the sum of WTP and QALY gains for all individuals. The other kind of studies under this category include those which infer the threshold value using the value of a statistical life (VSL), which can be computed through contingent valuation studies or revealed/stated preference methods that are commonly used in transport- and environment-related policy studies.[18]

Demand-side threshold estimates that are based on WTP surveys include several studies such as Shiroiwa et al.,[19] who conducted an international survey on general health to find that the average WTP for an additional QALY at a disaggregated level varied significantly from country to country, from £23,000 in the United Kingdom to US$62,000 in the United States and NT$2.1 million in Taiwan. Baker et al.,[20] based on aggregation of individual preferences for headache in England, found a threshold in the range of £22,570 to £41,350. Bobinac et al.[21] found a threshold for general health in the Netherlands in the range €80,800 to €113,000, by aggregation of individual preferences. Martín-Fernández et al.[22] estimated a range of WTP thresholds in Spain through disaggregated pooling of individual preferences, using different methods to elicit the WTP.

The other segment of literature on the demand-side includes Hirth et al.,[23] who determined the value of a QALY as implied by the VSL literature and compared this value with arbitrary thresholds for cost-effectiveness that have come into normal use. They identified 42 estimates of the value-of-life that were appropriate for inclusion through a literature search. Donaldson et al.[18] addressed the issue of threshold in two ways, first by modelling it using the United Kingdom health value of prevented fatality from the transport department to arrive at values in the range of £10,000 to £70,000 per QALY, and second by conducting a survey to test the feasibility of combining respondents' WTP and health state utility regressions. Via the survey, most methods of aggregating resulted in threshold

[*] Some recent papers from 2017-2018 were also added manually to our search.

values in the range of £18,000 to £40,000. These contrasted with the threshold range used by the National Institute for Health and Care Excellence (NICE) of £20,000 to £30,000.

A more detailed review of the demand-side estimates can be seen in Vallejo-Torres et al.[17] The multiple studies on the demand-side show considerable variation in estimates for the threshold, which can be attributed to the difference in methods of aggregation of WTP as well as the methods to make QALY adjustments.[24] VSL-based studies, on the other hand, can have estimates higher than that of WTP-based studies, but they too show variations depending on the method chosen to value the end of life.

Although the literature is comprised of mainly demand-side studies, many authors consider that, in the presence of a constrained budget for health, the shadow price (or opportunity cost) approach is better suited to elicit the threshold.[11, 25]

The supply-side studies are fewer in number, as this approach is heavily dependent on a continuous availability of complete information on the costs and QALYs gained for all possible interventions in the health care system. Moreover, the opportunity costs estimation requires the know-how of exactly which activities sector-wise will be displaced when a new health technology needs to be funded.[26, 27]

The supply-side estimates therefore can be inferred from past funding decisions,[27, 28, 29] or through empirical estimation of the marginal cost of a QALY.[11, 30] Using past funding decisions to infer the threshold is conceptually weak, on the grounds that current decisions of decision-makers may not always correlate with past decisions, and there are always multiple factors that affect a decision-making process that need complete transparency to ascertain.

A summary description of recent supply-side estimates of the threshold is provided in Table 1.

Threshold evaluation studies based on estimates of opportunity costs in the health care system reflected by the cost per QALY are now picking up momentum. Some of the earlier ones include Lichtenberg,[31] who developed a health production function using time series data in the United States from 1960 to 2001 to estimate a threshold of $11,000 per life-year (LY). In Spain, Puig-Junoy and Merino-Castello[32] applied a similar methodology using health spending and life expectancy at birth from 1960 to 1997, and estimated a cost per LY less than €13,000. These studies could not provide robust estimates of k, as it is difficult to disentangle impacts of time trends in expenditure from other temporal influences in health.

Woods et al.[15] have exploited the relationship between the GDP per capita of a country and the VSL to compute the cost-effectiveness threshold for several countries, from the lowest of $3 (in Malawi) to the highest of $8,018 (in Kazakhstan).

The above-mentioned studies are constrained by limitations owing to endogeneity between incomes/expenditures and health outcomes. There have been some recent studies that have tried to (partially) address this issue by adjusting the estimated impact on mortality to account for health-related quality of life (HRQoL) to estimate the marginal cost of a QALY, and at the same time using more sophisticated methods for treating endogeneity issues with instrumental variables and panel data sets. For example, Martin et al.[33, 34] and Claxton et al.[11] measured the cost per QALY using administrative data for primary care trusts (PCTs) in England using the spending data. Martin et al.[33, 34] used data for 2005/06 for five specific diseases, while Claxton et al.,[11] using expenditure data for 2008/09, provided an estimate for each of the 23 disease programs and combined the disease-specific values to arrive at a central estimate of £12,936 per QALY.

Lomas et al.[13] have extended the analysis by Claxton et al.[11] using mortality as the outcome variable and expenditure data as the explanatory variable, aggregated for 23 programs of health care along with additional non-clinical groups. They have created separate subgroups for "under-target" and "over-target" PCTs to compare how the expected opportunity costs of proposed health investment varies across the subgroups, thus throwing greater light on the budget impacts. Subgroup-specific elasticities are computed and the resulting threshold value is slightly lower than the one computed by Claxton et al.,[11] at £12,452 per QALY. This indicates that health opportunity costs can be underestimated for bigger investments if scale of the budget impact is not considered.

Edney et al.[35] have computed an estimate of the average opportunity cost to fund new health technologies in Australia over 2011/12 using instrumental variable two-stage least square regression analysis accounting for issues of endogeneity. Adapting on the methods of Claxton et al.,[11] they have estimated a reference ICER of AU$28,033, adding to the nascent and evolving literature on empirical estimates of threshold from the supply-side. Vallejo-Torres et al.[30] have also computed a similar estimate for Spain, using a panel data set across 17 regional health services from 2008 to 2012. Even though they use fixed effects estimation to address endogeneity, an instrumental variable (IV) estimation is also performed to fully capture all sources of variation within regions and years that correlate with expenditure and health outcomes. Their estimate varies between 21,000€ and 24,000€.

Summary

The literature on empirical estimates of cost-effectiveness thresholds across countries indicates that there are wide disparities in their values, owing to diverse assumptions placed in their estimation. Empirical estimates of thresholds differ not only with respect to the method of estimation (demand versus supply), but show variation also across different health care systems with different healthcare budgets. Demand-side thresholds are found to be on the higher end as of the current research, which can be explained to some extent through the theoretical underpinnings discussed in Section 2. It is still a nascent stage to comment on the exact linkages, as various assumptions used in empirical estimation need to be reevaluated. Linearity in the relation between health outcomes and expenditures is a core assumption of the supply-side analyses, which needs to be refined in the emerging research in this area. Since diminishing returns to health expenditures is theoretically established, it must be incorporated into empirical analysis as well. The decision to allocate scarce healthcare resources does need good supply-side threshold estimates, but, when the critical assumption of fixed budgets is relaxed, the support of demand-side thresholds is also relevant to throw light on the social preferences in deciding how to allocate any additional funds at the disposal of the health care system.

TABLE 1: Summary of the studies estimating relationship between health expenditures and health outcomes

#	Title	Authors	Year	Method of estimation	QALY adjustment	Disease area	Country/ Region	Range of threshold
1	Sources of U.S Longevity increase, 1960-2001	Lichtengber FR	2004	Time series: 1960-2001	NA	Total expenditure	United States	$11,000
2	Productividad marginal del gasto e innovacciones sanitarias	Puig-Junoy J, Merino Castelló A	2004	Time series: 1960-1997	NA	Total expenditure	Spain	€9,329-€11,076
3	Does health care spending improve health outcomes? Evidence from English programme budgeting data	Martin S, Rice N, Smith PC	2008	Instrumental variable	Utility scores by ICD-10 codes from HODaR project	Cancer, circulatory diseases	England	£19,070, £11,960
4	Comparing costs and outcomes across programmes of health care	Martin S, Rice N, Smith P	2012	Instrumental variable	Utility scores by ICD-10 codes from HODaR project	Cardiovascular, respiratory, gastrointestinal, diabetes	England	£12,593, £13,256, £30,400, £47,069
5	Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold	Claxton K, Martin S, Soares M, et al.	2015	Instrumental variable	Utility scores by ICD-10 codes from HODaR project	23 program budget categories	England	£12,936
6	Country-level cost-effectiveness thresholds: Initial estimates and need for further research	Woods B, Revill P, Sculpher M, Claxton K	2016	Income elasticities of VSL	NA	Overall	Malawi, Cambodia, El Salvador, Kazakhstan	$3-$116, $44-$518, $422-$1,967, $4,485-$8,018
7	Estimating cost-	Vallejo-Torres L,	2017	Panel fixed effects -	Adjusted for	Total	Spain	€21,000-€24,000

#	Title	Authors	Year	Method of estimation	QALY adjustment	Disease area	Country/ Region	Range of threshold
	effectiveness threshold for Spanish NHS	Garcia-Lorenzo B, Serrano-Aguillar P		Instrumental variable	HRQoL by using EQ-5D weights	expenditure		
8	Estimating the reference incremental cost-effectiveness ratio for the Australian health system	Edney LC, Haji Ali Afzali H, Cheng TC, Karnon J	2017	Instrumental variable - 2 SLS	Utility scores by SF-36 data and SF-6D algorithm	Total expenditure	Australia	AU$28,033
9	Resolving the "cost-effectiveness but unaffordable" paradox: Estimating health opportunity costs to nonmarginal budget impacts	Lomas J, Claxton K, Martin S, Soares M	2018	Instrumental variable	Utility scores by ICD-10 codes from HODaR project	23 program budget categories	England	£12,452 (non-marginal budget impact)

SECTION 4: Incorporating equity weights into consideration of value-based pricing: evidence and issues

The central tenet of value-based pricing is that the price paid should reflect the value of its health and resource impacts. In the context of the supply-side cost-effectiveness threshold model, this means that, when a technology has a positive budget impact, the value of the health care displaced to cover that budget impact must be less than the value of the additional health produced. In the context of the demand-side cost-effectiveness threshold, the value-based price is the price that produces an ICER that is exactly equal to the maximum WTP for health.[36, 37]

Many of the criticisms of value-based pricing are concerned with the specification of the value of the health produced.[38] Few, if any, health economists would argue that the value that individuals attach to health gains should be independent of the characteristics of the individuals to whom the gains accrue, or of the characteristics of the target health condition.[18, 39, 40] However, conceptual agreement on this point appears to be the limit of consensus in public debates regarding the "value" in value-based pricing. Paulden et al.,[41] reporting a scoping review of value arguments in the context of orphan drug reimbursement, identified 19 candidate determinants of value from 43 studies; the number of papers citing any one of these candidates ranged from one to 23.

Paulden et al.[41] also described a conceptual model for the incorporation of these broader measures of value into reimbursement decisions, within the supply-side threshold model. The implications of this model were elucidated further in subsequent publications. The central insight from this work is that, whatever definition of value is used to assess a technology, horizontal equity requires that the same definition is used to characterize the opportunity cost (value foregone to fund the additional cost of the new technology).

To recap, horizontal equity requires that "equals are treated equally," while vertical equity requires that "unequals are treated unequally." If a decision-maker chooses to value all health equally, irrespective of the characteristics of the individual who receives or loses it – frequently called the "QALY is a QALY" position[42, 43] –, this is a specific vertical equity position. Its advantage is that a decision-maker can apply this to the evaluation of a technology without having to know anything about the characteristics of the individuals who bear the opportunity cost of adopting that technology. A departure from this vertical equity position, for example to take account of the age of beneficiary of the technology (perhaps children or the elderly), complicates the decision-maker's task. They must know which patients bear the opportunity cost and hence to what degree the additional value applies, for example, how many children lose out. Failure to do this breaches the requirement for horizontal equity, as the decision-maker can no longer be confident that individuals with the same characteristics are being treated equally in their decision-making.[44, 45]

There is a small but growing body of empirical research on the values that societies should consider in healthcare reimbursement. Following on from McCabe et al.'s[46] call for empirical research on the question, in the context of evaluating ultra-orphan drugs, Desser and colleagues[47, 48, 49] produces a series of studies on the characteristics that might be taken into account, in a Norwegian context. These have been followed up with studies by Linley and Hughes in the United Kingdom,[50] and Shah et al.,[51] Rowen et al.,[52] and Chim et al.[53] in Australia. To date, the research overwhelming confirms that the greatest value is attached to the magnitude of the health gain. Many characteristics that have been argued for in the policy literature – such as prevalence and proximity to the end of life, and an "innovation" premium – have been consistently rejected. A small number of characteristics do consistently receive support including the severity of illness: a) health gained by individuals with more severe illness being valued more highly than health gained by those with mild illness; b) treatments for illnesses for which there is no alternative therapy; and c) very high cost treatments, interpreted as treatments with catastrophic costs, that is, those that exceed the disposable income of the household.

Wailoo et al.[54] consider the question of the how evidence on the value of additional characteristics should be used by decision-makers. They observe that these additional characteristics would need to be incorporated into the value function, similar to the current utility algorithms that are used to calculate utilities from health state data. On this basis, they argue that the questions identified by Dolan[55] in the context of the measurement and valuation of health would still need to be considered. Hence, having established what is valued (for example, health plus the severity of the target condition and the [non-]existence of alternative therapies), decisions about whose value of these characteristics should be used (for example, general population or patients), how those values should be obtained (for example, standard gamble, time trade off, or discrete choice experiment), and how values for different combinations of these characteristics should be estimated (for example, linear additive or multiplicative functions) would have to be made. Some nine years later, there is no research that answers these questions. As Wailoo and colleagues discussed, given the current evidence, the application of empirical equity weights to inform value assessment will have to wait.

Summary

While quantitative incorporation of equity considerations into cost-effectiveness analyses is not currently possible, the current evidence does support the incorporation of at least some additional value characteristics, in considerations of the value of new drugs. These considerations must of necessity be dealt with through qualitative consideration within the decision-making process. However, these considerations should be incorporated into the decision-making process in a way that is cognizant of the need to respect both horizontal and vertical equity.[10, 56]

SECTION 5: Exposition of mechanisms for incorporating societal contribution to global pharmaceutical research and development into value-based pricing assessments

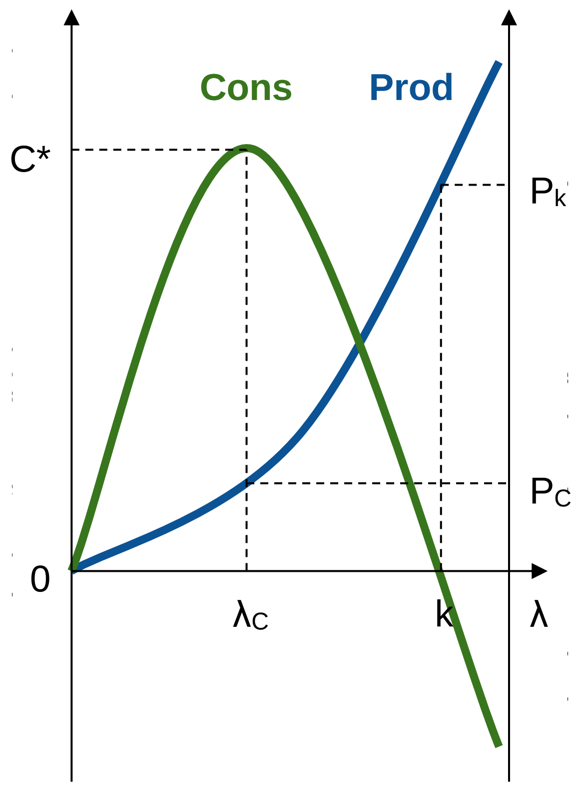
The existing supply- and demand-side approaches to estimating thresholds, as summarized in Section 3, do not take into account the allocation of consumer and producer surplus. Existing models also do not take into account strategic pricing behaviour on the part of manufacturers that might be expected to follow any specification of a threshold in practice.

The draft paper attached to this report proposes a new conceptual model of the threshold that incorporates both of these considerations. This model builds upon existing supply- and demand-side approaches to the threshold and integrates considerations from both. It reveals that the appropriate threshold depends upon a number of factors, including:

1. the conventional supply-side threshold (that is, the shadow price of the healthcare budget, and the subject of recent empirical work described in Section 3);
2. the conventional demand-side threshold (that is, the monetary value of a unit of benefit produced by the health care system, also the subject of empirical work);
3. the policy objective, specifically the desired allocation of the total surplus from the adoption of new technologies between consumers (patients) and producers (the manufacturers of new technologies adopted by the health care system); and
4. the distribution of reserve prices (and hence the distribution of reserve ICERs) for new technologies; this distribution reflects the minimum ICER at which manufacturers are willing to supply each new technology to the health care system, given the costs of production and the desire to make an acceptable return on the manufacturer's investment in research and development (R&D).

The attached draft paper describes in detail the assumptions made in the proposed model, and how the supply- and demand-side approaches can be integrated. The key findings from this draft paper are reproduced below in Figure 2 and Table 2.

FIGURE 2: Consumer and producer threshold curves, reflecting the relationship between the threshold (λ), net population benefit (consumer surplus), and manufacturer profit (producer surplus)



Where the objective is to maximize consumer surplus, the optimal threshold is λ_C. Note that λ_C is lower than *k*, the conventional supply-side threshold, which in turn may be expected to be lower than *v*, the conventional demand-side threshold (see Section 2). Thus the proposed model finds that, under this objective, the specified threshold should be lower than that implied by both conventional approaches. If, in addition, there is a desire that producer surplus comprise a guaranteed proportion of the combined surplus, this may require that the threshold be increased above λ_C but no higher than *k*, until this proportion is reached.

Where the objective is to maximize producer surplus, the optimal threshold is infinitely high. However, this results in negative consumer surplus; if there is also a desire for consumer surplus to be non-negative, then the optimal threshold is *k*. If, in addition, there is a desire that consumer surplus comprise a guaranteed proportion of the combined surplus, this may require that the threshold be lowered below *k* until this proportion is reached.

Finally, where the objective is to maximize the combined surplus, the optimal threshold lies somewhere above λ_C, with its precise location dependent upon the shape of each threshold curve and the conversion rate between consumer and producer surplus. If there is also a desire that both consumer and producer surplus be non-negative, the optimal threshold lies somewhere above λ_C, but no higher than *k*.

Based on recent empirical estimates of supply-side thresholds (£12,936, €24,870, and AU$28,033 per QALY in England, Spain, and Australia, respectively), the proposed model implies that, if decision-makers in these countries have a primary concern for maximizing consumer surplus, then thresholds lower than these should be specified in practice. The use of higher thresholds is consistent with an objective of maximizing producer surplus, subject to a weak concern for consumer surplus that serves only to limit the extent to which it is negative.

Specifying a threshold in Canada requires consideration of all of the factors described above. Particular attention should be paid to the policy objective. If Canadian decision-makers believe that new technologies should be adopted only if they provide positive consumer surplus, and if the assumptions of the proposed model are considered to be applicable (including strategic pricing on the part of manufacturers), then a threshold should be specified that is no higher than the shadow price of the healthcare budget, as estimated through a conventional supply-side approach. If there is a concern that consumer surplus be maximized, then a lower threshold should be adopted, but estimating precisely what this threshold is requires novel empirical research into the distribution of reserve ICERs across new technologies. Alternatively, if there is a concern that manufacturers receive a guaranteed proportion of the total surplus from new technologies, then a threshold somewhere between the two should be considered. This final concern might be borne out of a view that Canada should contribute a certain share towards global pharmaceutical R&D. It should, however, be considered that increases in the threshold to provide this share come at the expense of consumer surplus, that is, the benefit to patients derived from the health care system.

TABLE 2: Optimal threshold (λ*) or range containing optimal threshold, for each objective

Policy objective	Optimal threshold or range containing optimal threshold	Comments
Maximize consumer surplus	$\lambda^* = \lambda_C$	Consumer surplus is maximized by specifying a threshold of λ_C.
Maximize consumer surplus, subject to producer surplus comprising a guaranteed proportion of the combined surplus	$\lambda_C \leq \lambda^* \leq k$	The proportion of the combined surplus allocated to producers increases above λ_C. If producer surplus comprises the required proportion at λ_C, then this is the optimal threshold. If not, the threshold should be progressively increased until the required proportion is achieved.
Maximize producer surplus	$\lambda^* = \infty$	Producer surplus is maximized with an infinitely high threshold.
Maximize producer surplus, subject to consumer surplus being non-negative	$\lambda^* = k$	Since producer surplus increases with the threshold, and consumer surplus is negative at any threshold above k, this objective is satisfied by specifying a threshold of k.
Maximize producer surplus, subject to consumer surplus comprising a guaranteed proportion of the combined surplus	$0 \leq \lambda^* \leq k$	The maximum threshold at which each is non-negative is k. The optimal threshold is derived by progressively lowering the threshold from k until the required proportion of consumer surplus is achieved.
Maximize the combined surplus	$\lambda_C < \lambda^* \leq \infty$	Consumer and producer surplus both increase with the threshold up to λ_C. Above λ_C, consumer surplus falls and producer surplus increases. The optimal threshold depends upon the shape of each threshold curve but must exceed λ_C.
Maximize the combined surplus, subject to consumer surplus being non-negative	$\lambda_C < \lambda^* \leq k$	Since consumer and producer surplus both increase with the threshold up to λ_C but consumer surplus is negative above k, the optimal threshold must lie between λ_C and k.

SECTION 6: Summary

In this report, we have described the relationship between demand- and supply-side approaches to the cost-effectiveness threshold, and have identified several reasons why the demand-side threshold might be higher than the supply-side threshold in practice.

In Section 3, we identified three distinct approaches for empirically estimating the cost-effectiveness threshold. We identified a number of estimates of willingness-to-pay (WTP) for health (demand-side) for different health care systems, and observed potentially important differences in the reported values according to the method used and the health care system context. A small number of studies reported indirect estimates of WTP based upon statistical analysis of observed funding decisions by healthcare payers. It is not possible to say whether these estimates are examples of a demand- or supply-side threshold, as decision-makers are not explicit about which model they are operating under. Further, these studies are subject to the concerns of unknown confounders that affect all observation analyses. A small number of studies reported supply-side estimates of the threshold. The United Kingdom estimates have used aggregate level data, but more recent estimates for Spain and Australia are based upon individual patient-level data. These supply-side estimates are consistently lower than published demand-side estimates and also conventional thresholds values assumed by decision-makers in the absence of empirical estimates. We did not identify empirical estimates of supply-side thresholds for Canada as a whole or for individual Canadian provinces.

In Section 4, we considered the incorporation of wider value considerations into cost-effectiveness-based health technology assessment processes. We identified a range of additional value characteristics being proposed by different authors. However, there was no clear consensus across authors about which of the many values should be considered. An emerging empirical evidence base was also identified on which value considerations matter to the general public. A societal preference for placing additional value on treatments for severe conditions, conditions for which there is no current therapy, and high cost/catastrophic cost treatments appeared to be consistent across studies. We identified a number of substantive challenges to quantitative approaches to incorporating wider value considerations into decision-making processes. First, there was a limited evidence base to support the choice of which wider value considerations to include. Second, there is even less evidence of the appropriate weight to attach to each consideration, relative to each other and to health outcomes. Finally, there is no evidence on the appropriate functional form for combining the value attached to health and each additional consideration. A further substantive challenge, specific to decisions operating under the supply-side threshold, relates to ensuring adherence with horizontal and vertical equity requirements. Unless decision-makers know the characteristics of the individuals who bear the opportunity cost of adopting a new technology due to its positive budget impact, as well as the characteristics of the beneficiaries of the new technology, a reimbursement decision based on wider value considerations may displace more highly-valued care than it produces; this is because the wider value characteristics may be more prevalent amongst those who bear the opportunity cost.

In Section 5, we reported a *de novo* conceptual model of the use of cost-effectiveness thresholds to make reimbursement decisions, examining a number of policy objectives regarding the distribution of consumer surplus (net population benefit for patients) and producer surplus (manufacturer profit). This novel framework is motivated from the observation that the cost-effectiveness threshold can operate as a signal to investors in developing innovative technologies, but is also subject to strategic pricing behaviour on the part of manufacturers. A summary of the findings of this model was provided in the previous section, and a draft paper providing more detail is also attached to this report. The proposed model implies that the optimal threshold depends on the policy objective, and is generally lower than *k* (the threshold from the conventional supply-side model) if improving consumer surplus is a policy concern. The model incorporates two threshold curves, reflecting the relationship between the threshold and consumer and producer surplus, respectively. The shape of the consumer threshold curve is influenced by strategic behaviour on the part of manufacturers. The share to society reflecting a target return on the societal investment in a) basic research and b) healthcare infrastructure is required to realize the value of the technology. Producer surplus provides a reward to manufacturers for their R&D in excess of the price that drives their reserve ICER, which incorporates their target return on R&D investment. The additional return on investment represents supra-normal profits, which act as a signal to other investors to enter into the market. Additional investors should, in principle, increase the probability of R&D investments leading to technologies that meet currently unmet needs.

The work presented in this report provides a number of key insights to inform policy debates around the move towards pharmaceutical price setting based upon cost-effectiveness thresholds: 1) demand-side thresholds will, except under unusual circumstances, be higher than supply-side thresholds; 2) there is a need for empirical research on both demand- and supply-side thresholds in the Canadian context; 3) wider value considerations, sometimes called *equity* or *ethical concerns*, must remain a qualitative process given the current evidence base – further research on which value considerations to include and how much value to attach to each (compared to health and each other) would be valuable; 4) specifying a threshold provides a mechanism for allocating the total surplus from new technologies between consumers and producers; and 5) where manufacturers strategically price to the threshold, specifying a cost-effectiveness threshold lower than *k* is required if the adoption of new technologies is to increase consumer surplus – the supra-normal profits that arise from this behaviour also provide an incentive for R&D, particularly into technologies that may provide substantial benefit to the health care system at low cost, where the potential for supra-normal profits is maximized.

Policy considerations

Value-based pricing has typically been used in reimbursement decision-making processes rather than a free-standing price setting activity. The use of the supply-side cost-effectiveness threshold to identify a maximum price for a technology is equivalent to defining an excessive price where the expected benefit displaced through its incremental cost is greater than the benefit the technology is expected to produce.

Identifying the cost-effectiveness threshold to be used in the price setting process can, conceptually, be undertaken using either a social WTP approach (the demand-side model) or a health benefit opportunity cost approach (the supply-side model). It is unlikely that the demand-side threshold (*v*) and the supply-side threshold (*k*) will be equal. Health care in Canada is, primarily, provided through publicly funded health care systems. There is increasing political pressure to bend the cost curve in Canadian health care, and professional organizations are advocating for clinical practice that promotes the financial sustainability of health care systems. In this policy context, it seems appropriate to design policy around the assumption that healthcare budgets are constrained, and hence the supply-side model is an appropriate framework for identifying the maximum cost-effectiveness threshold (*k*) for a value-based pricing policy.

The role of demand-side estimates of the WTP for health (v) is twofold. First, these estimates should inform the size of the healthcare budget; second, it can be used as the exchange rate between producer and consumer surplus, which would be essential if the threshold (λ) was to be set on the basis of sharing the total value produced by adopting new technologies between producers and consumers.

To identify the appropriate value of λ to use in a value-based pricing policy, decision-makers need to be clear about the policy objectives, specifically about how the additional welfare created by innovative new technologies should be shared between producers (manufacturers) and consumers (healthcare payers). Setting λ equal to k is consistent with the objective of maximizing producer surplus, subject to the constraint that consumer surplus is non-negative. Maximizing consumer surplus from new technologies requires that λ be set considerably lower than k. Values of λ beyond λ_C reduce the impact of new technologies on total population health benefit but may well be justified for equity reasons, for example, if the marginal technologies adopted address significant unmet needs or help otherwise disadvantaged members of society. When all manufacturers are able to price up to the threshold, the lost health benefit of setting λ above λ_C will be greater than if higher values of λ are only applied to a subgroup of technologies that meet the equity criteria. There is an emerging empirical evidence base on societal preferences for equity considerations, which can be used to guide decision-makers' judgements. That said, the maturity of this evidence is not sufficient for the considerations to be included in quantitative analyses.

Values of λ beyond k are coherent with a policy objective that gives primacy to producer surplus, as they entail sacrificing current total population health in order to increase producer surplus. Values of λ below all the expected reserve ICERs would mean that no new technologies will be available. The implication of this is that the health care system is not interested in improvements in efficiency through the adoption of new technologies.

The ability of decision-makers to make judgements about reserve ICERs, and hence the impact of any specific value of λ on access to new therapies, is hampered by the information asymmetry between producers and payers with regard to R&D and the upstream and downstream cost of goods. While some legislators, particularly in the United States, are examining mandating the disclosure of R&D costs within price negotiations, this data is not typically accessible outside of the companies. In these circumstances, setting λ at any specific level will reflect a searching function, whereby the responses of manufacturers provide indirect information on their reserve ICERs. However, this information is also subject to strategic behaviour by the manufacturers.

In summary, the current model of the supply-side cost-effectiveness threshold incorrectly suggests that total population health benefit is maximized by setting λ equal to k. This condition actually maximizes producer surplus, subject to the constraint that consumer surplus is non-negative. Value-based pricing using the supply-side cost-effectiveness threshold is feasible as a mechanism for operationalizing the concept of an excessive price. It is also possible to incorporate equity considerations into such a framework. Empirical evidence on a) the value of k for Canadian healthcare payers, b) the reserve prices and hence ICERs for Canadian pharmaceutical products, and c) the demand-side cost-effectiveness threshold (WTP for health) of Canadian citizens would significantly enrich the value of implementing this approach.

References

1. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in health and medicine*. Oxford University Press; 1996.
2. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. Oxford University Press; 2015.
3. Weinstein M, Zeckhauser R. Critical ratios and efficient allocation. *J Public Econ* 1973;2:147-157.
4. Azimi NA, Welch HG. The effectiveness of cost-effectiveness analysis in containing costs. *J Gen Intern Med* 1998;13:664-669.
5. Garber AM, Phelps CE. Economic foundations of cost-effectiveness analysis. *J Health Econ* 1997;16:1-31.
6. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: What it is and what that means. *Pharmacoeconomics* 2008;26:733-744.
7. Eichler H-G, Kong SX, Gerth WC, Mavros P, Jönsson B. Use of cost-effectiveness analysis in health-care resource allocation decision-making: how are cost-effectiveness thresholds expected to emerge? Value Health. 2004;7: 518–528.
8. Marseille E, Larson B, Kazi DS, Kahn JG, Rosen S. Thresholds for the cost-effectiveness of interventions: Alternative approaches. *Bull World Health Organ* 2014;93:118-124.
9. Guilbert JJ. The world health report 2002 - reducing risks, promoting healthy life. *Educ Health* 2003;16:230.
10. Culyer AJ. Cost-effectiveness thresholds in health care: A bookshelf guide to their meaning and use. *Health Econ Policy Law* 2016;11:415-432.
11. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, et al. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technol Assess* 2015;19:1-503, v-vi.
12. Paulden M, O'Mahony J, McCabe C. Determinants of change in the cost-effectiveness threshold. *Med Decis Making* 2017;37:264-276.
13. Lomas J, Claxton K, Martin S, Soares M. Resolving the "cost-effective but unaffordable" paradox: Estimating the health opportunity costs of nonmarginal budget impacts. *Value Health* 2018; doi:10.1016/j.jval.2017.10.006.
14. Culyer A, McCabe C, Briggs A, Claxton K, Buxton M, Akehurst R, et al. Searching for a threshold, not setting one: The role of the National Institute for Health and Clinical Excellence. *J Health Serv Res Policy* 2007;12:56-58.
15. Woods B, Revill P, Sculpher M, Claxton K. Country-level cost-effectiveness thresholds: Initial estimates and the need for further research. *Value Health* 2016;19:929-935.
16. Baker R, Chilton S, Donaldson C, Jones-Lee M, Lancsar E, Mason H, et al. Searchers vs surveyors in estimating the monetary value of a QALY: Resolving a nasty dilemma for NICE. *Health Econ Policy Law* 2011;6:435-447.

17. Vallejo-Torres L, García-Lorenzo B, Castilla I, Valcárcel-Nazco C, García-Pérez L, Linertová R, et al. On the estimation of the cost-effectiveness threshold: Why, what, how? *Value Health* 2016;19:558-566.

18. Donaldson C, Baker R, Mason H, Jones-Lee M, Lancsar E, Wildman J, et al. The social value of a QALY: Raising the bar or barring the raise? *BMC Health Serv Res* 2011;11:8.

19. Shiroiwa T, Sung Y-K, Fukuda T, Lang H-C, Bae S-C, Tsutani K. International survey on willingness-to-pay (WTP) for one additional QALY gained: What is the threshold of cost effectiveness? *Health Econ* 2010;19:422-437.

20. Baker R, Bateman I, Donaldson C, Jones-Lee M, Lancsar E, Loomes G, et al. Weighting and valuing quality-adjusted life-years using stated preference methods: Preliminary results from the Social Value of a QALY Project. *Health Technol Assess* 2010;14:1-162.

21. Bobinac A, van Exel J, Rutten FFH, Brouwer WBF. The value of a QALY: Individual willingness to pay for health gains under risk. *Pharmacoeconomics* 2014;32:75-86.

22. Martín-Fernández J, Polentinos-Castro E, del Cura-González MI, Ariza-Cardiel G, Abraira V, Gil-LaCruz AI, et al. Willingness to pay for a quality-adjusted life year: An evaluation of attitudes towards risk and preferences. *BMC Health Serv Res* 2014;14:287.

23. Hirth RA, Chernew ME, Miller E, Fendrick AM, Weissert WG. Willingness to pay for a quality-adjusted life year: In search of a standard. *Med Decis Making* 2000;20:332-342.

24. Gyrd-Hansen D. Willingness to pay for a QALY. *Health Econ* 2003;12:1049-1060.

25. Karlsberg Schaffer S, Cubi-Molla P, Devlin N, Towse A. *Shaping the research agenda to estimate relevant cost-effectiveness thresholds for health technology assessment decision-making: Report for ABPI*. London, The Office of Health Economics (OHE); 2016.

26. Towse A. Should NICE's threshold range for cost per QALY be raised? Yes. *BMJ* 2009;338:b181.

27. Appleby J, Devlin N, Parkin D, Buxton M, Chalkidou K. Searching for cost effectiveness thresholds in the NHS. *Health Policy* 2009;91:239-245.

28. Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Econ* 2004;13:437-452.

29. Schilling C, Mortimer D, Dalziel K. Using CART to identify thresholds and hierarchies in the determinants of funding decisions. *Med Decis Making* 2017;37:173-182.

30. Vallejo-Torres L, García-Lorenzo B, Serrano-Aguilar P. Estimating a cost-effectiveness threshold for the Spanish NHS. *Health Econ* 2017; doi:10.1002/hec.3633.

31. Lichtenberg FR. Sources of U.S. longevity increase, 1960–2001. *Q Rev Econ Finance* 2004;44:369-389.

32. Puig-Junoy J, Merino-Castelló A. Productividad marginal del gasto e innovaciones sanitarias. Resultados empíricos y lecciones para España. *Más recursos para la salud* 2004;133-154.

33. Martin S, Rice N, Smith PC. Does health care spending improve health outcomes? Evidence from English programme budgeting data. *J Health Econ* 2008;27:826-842.

34. Martin S, Rice N, Smith PC. Comparing costs and outcomes across programmes of health care. *Health Econ* 2012;21:316-337.

35. Edney LC, Haji Ali Afzali H, Cheng TC, Karnon J. Estimating the reference incremental cost-effectiveness ratio for the Australian health system. *Pharmacoeconomics* 2017; doi:10.1007/s40273-017-0585-2.

36. Towse A. Value based pricing, research and development, and patient access schemes. Will the United Kingdom get it right or wrong? *Br J Clin Pharmacol* 2010;70:360-366.

37. Claxton K. OFT, VBP: QED? *Health Econ* 2007;16:545-558.

38. Sussex J, Towse A, Devlin N. Operationalizing value-based pricing of medicines: A taxonomy of approaches. *Pharmacoeconomics* 2013;31:1-10.

39. Culyer AJ. The morality of efficiency in health care—some uncomfortable implications. *Health Econ* 1992;1:7-18.

40. Culyer AJ, Bombard Y. An equity framework for health technology assessments. *Med Decis Making* 2012;32:428-441.

41. Paulden M, Stafinski T, Menon D, McCabe C. Value-based reimbursement decisions for orphan drugs: A scoping review and decision framework. *Pharmacoeconomics* 2015;33:255-269.

42. Donaldson C, Atkinson A, Bond J, Wright K. Should QALYs be programme-specific? *J Health Econ* 1988;7:239-257.

43. Weinstein MC. A QALY is a QALY is a QALY — Or is it? *J Health Econ* 1988;7:289-290.

44. Paulden M, O'Mahony JF, Culyer AJ, McCabe C. Some inconsistencies in NICE's consideration of social values. *Pharmacoeconomics* 2014;32:1043-1053.

45. Paulden M, O'Mahony JF, Culyer AJ, McCabe C. Objectivity and equity: Clarity required. A response to Hill and Olson. *Pharmacoeconomics* 2014;32:1249-1250.

46. McCabe C, Claxton K, Tsuchiya A. Orphan drugs and the NHS: Should we value rarity? *BMJ* 2005;331:1016-1019.

47. Desser AS, Gyrd-Hansen D, Olsen JA, Grepperud S, Kristiansen IS. Societal views on orphan drugs: Cross sectional survey of Norwegians aged 40 to 67. *BMJ* 2010;341:c4715.

48. Desser AS, Olsen JA, Grepperud S. Eliciting preferences for prioritizing treatment of rare diseases: The role of opportunity costs and framing effects. *Pharmacoeconomics* 2013;31:1051-1061.

49. Desser AS. Prioritizing treatment of rare diseases: A survey of preferences of Norwegian doctors. *Soc Sci Med* 2013;94:56-62.

50. Linley WG, Hughes DA. Societal views on NICE, cancer drugs fund and value-based pricing criteria for prioritising medicines: A cross-sectional survey of 4118 adults in Great Britain. *Health Econ* 2013;22:948-964.

51. Shah KK, Tsuchiya A, Wailoo AJ. Valuing health at the end of life: A stated preference discrete choice experiment. *Soc Sci Med* 2015;124:48-56.

52. Rowen D, Brazier J, Mukuria C, Keetharuth A, Risa Hole A, Tsuchiya A, et al. Eliciting societal preferences for weighting QALYs for burden of illness and end of life. *Med Decis Making* 2016;36:210-222.

53. Chim L, Salkeld G, Kelly P, Lipworth W, Hughes DA, Stockler MR. Societal perspective on access to publicly subsidised medicines: A cross sectional survey of 3080 adults in Australia. *PLoS One* 2017;12:e0172971.

54. Wailoo A, Tsuchiya A, McCabe C. Weighting must wait. *Pharmacoeconomics* 2009;27:983-989.

55. Dolan P. Chapter 32: The measurement of health-related quality of life for use in resource allocation decisions in health care. In: Culyer AJ, Newhouse JP, editors. *Handbook of health economics*. Elsevier; 2000. p. 1723-1760.

56. Paulden M. Recent amendments to NICE's value-based assessment of health technologies: implicitly inequitable? *Expert Rev Pharmacoecon Outcomes Res* 2017;17:239-242.

Appendix 1: Strategic Behaviour and the Cost-Effectiveness Threshold: A New Conceptual Model

Dr Mike Paulden, PhD

Contact details

Dr Mike Paulden, PhD
Assistant Professor, School of Public Health
University of Alberta
3-264 ECHA, 11405 87 Ave NW
Edmonton, AB, T6G 1C9
CANADA

Tel: +1 (780) 492-9921
Email: paulden@ualberta.ca

Classification codes

I110, I180, H43

Funding sources

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

Acknowledgements

The author is grateful to Christopher McCabe for suggestions made on an earlier draft. All mistakes remain the responsibility of the author.

Keywords

Cost-effectiveness analysis, cost-effectiveness threshold, decision making, health technology assessment

Introduction

Many publicly funded health care systems use ‘health technology assessment’ (HTA) to inform decisions regarding which new technologies to fund. An important component of such assessments is a determination of which new technologies are “cost-effective”. This typically involves a comparison of the incremental cost-effectiveness ratio (ICER) of each new technology to a “cost-effectiveness threshold”.

Recent years have seen a number of advancements in our theoretical and empirical understanding of how the cost-effectiveness threshold should be specified (Vallejo-Torres et al. 2016; Paulden et al. 2017; Lomas et al. 2018). It is now broadly accepted that there are two conceptually different theoretical approaches to specifying such thresholds - often characterised as ‘supply-side’ and ‘demand-side’ approaches - and that the relevance of each depends upon the context in which decisions are made.

A conventional supply-side approach for specifying the threshold is widely considered to be appropriate for decisions regarding whether or not to adopt new health technologies into a health care system subject to a constrained budget, where the opportunity cost is expected to fall elsewhere within the same budget, and where the policy concern is to maximize some measure of ‘benefit’ across the population as a whole (Claxton et al. 2011). This ‘benefit’ is often assumed to be measurable in quality-adjusted life years (QALYs), although the implications of the supply-side model also hold if other measures of benefit are used instead.

By contrast, a conventional demand-side approach for specifying the threshold is considered to be appropriate where the opportunity cost of adopting new technologies is expected to fall upon individual consumption and the policy concern is whether or not the consumption value of any health gains exceeds the associated consumption loss.

Regardless of the approach taken, specifying thresholds in practice requires empirical evidence. Although empirical estimates of demand-side thresholds have been available for many decades, it is only over recent years that estimates of supply-side thresholds have been published (Vallejo-Torres et al. 2016). Recent empirical studies of public health care systems in England, Spain and Australia have reported base-case estimates of supply-side thresholds of £12,936, €24,870 and AU$28,033 per QALY, respectively (Claxton et al. 2015; Vallejo-Torres et al. 2017; Edney et al. 2017).

The conventional supply-side approach

The recently published empirical estimates of supply-side thresholds reflect the relationship between marginal changes in health expenditure and health outcomes within each health care system. This relationship is frequently referred to as the 'shadow price' of the health care system budget constraint. For example, the AU$28,033 per QALY estimate from the recent Australian study represents the shadow price of the Australian health care system budget, reflecting the relationship between marginal health expenditures and health outcomes. This means that if expenditures on health care were to be increased by AU$28,033, all other things equal, we would expect one additional QALY to be produced. Conversely, if expenditures were to be reduced by AU$28,033 then we would expect one fewer QALY to be produced. Since, under the conventional supply-side model of the threshold, the health care budget is assumed to be constrained, every AU$28,033 spent on a *new* health technology would be expected to displace AU$28,033 from existing health care services, and hence would be expected to displace one QALY.

Proponents of a supply-side approach conventionally advocate for using this shadow price directly as the cost-effectiveness 'threshold' to which new technologies are compared. For example, if adopting a new technology would cost an additional AU$100,000 and provide two additional QALYs, then the technology's 'incremental cost-effectiveness ratio' (ICER) of AU$50,000 per QALY would be compared directly to the AU$28,033 estimate of the shadow price; since the ICER is higher than the shadow price, the technology would not be considered 'cost-effective'. If, instead, the technology provided 5 additional QALYs, then the ICER would be AU$20,000 per QALY; since this is lower than the shadow price, the technology would be considered cost-effective.

This conventional approach to specifying a supply-side threshold is consistent with a policy objective under which new technologies are considered cost-effective if, and *only if*, the benefits they provide exceed the benefits they displace. For example, if the shadow price is AU$28,033 per QALY, then a new technology that costs an additional AU$100,000 would be expected to displace approximately 3.6 QALYs. If the new technology provides more than 3.6 additional QALYs, its ICER will be lower than AU$28,033 per QALY and so it would be considered cost-effective. Conversely, if the technology provides fewer than 3.6 additional QALYs, its ICER will be higher than AU$28,033 per QALY and so it would *not* be considered cost-effective.

Limitations of existing approaches

Many authors consider the conventional supply-side approach to be consistent with an objective of maximizing total benefits across the population (Claxton et al. 2011; Revill et al. 2015; Remme et al. 2017; Olsen 2017). However, this is mistaken for at least two reasons.

First, as Pekarsky has demonstrated, if the health care system is *inefficient* then there are at least two different ‘shadow prices’ to consider: the relationship between marginal *reductions* in expenditure and health outcomes associated with a contraction of existing health services through displacement (as considered in the conventional supply-side model), and the relationship between marginal *increases* in expenditure and health outcomes associated with a potential expansion of other health care services (Pekarsky 2012). In an efficient health care system these two shadow prices are identical, so estimating only one of these shadow prices is sufficient, but in an inefficient health care system these shadow prices will diverge, with potentially important policy implications. If the policy objective is simply to ensure that new technologies do not displace more benefits than they provide, then there is no need to estimate this second shadow price - policy makers need only compare the benefits from the new technology to the benefits forgone through the displacement of other health services (Paulden et al. 2014). However, if the objective is to maximize total benefits across the population, then both shadow prices must be considered (Eckermann & Pekarsky 2014). This is because the opportunity cost of adopting a new technology is not merely the benefits forgone through displacement, but also the forgone opportunity to use the resources freed up through displacement for the expansion of other (more efficient) health care services. The threshold to which new technologies must be compared in order to take this opportunity cost into account is lower than the first of these shadow prices, implying that a threshold set in accordance with the conventional supply-side approach would be too high.

Second, the use of any cost-effectiveness threshold in practice inevitably leads to strategic behaviour by the manufacturers of health technologies. This includes the practice of ‘pricing to the threshold’, under which a manufacturer will set the price of a new technology such that the ICER falls marginally below, but as close as possible to, the specified threshold. Alternatively, a manufacturer may initially price a technology such that the ICER is higher than the specified threshold, with the intention of later negotiating the price down with payers; the price may then be expected to fall through negotiations until the ICER reaches the threshold. Regardless of the mechanism used, there is clearly a strong incentive for each manufacturer to behave in a way that results in a final ICER (following pricing and negotiations) as close as possible to the specified threshold; if they did not, and the final ICER was substantially below the threshold, then the manufacturer would be unnecessarily foregoing potential profits. Under a conventional supply-side approach, where the threshold determined directly be the first shadow price described above, the result of this strategic pricing behaviour by manufacturers is that the expected benefits from new technologies are exactly offset by the benefits expected to be displaced. For example, if the threshold is set at AU$28,033 per QALY, and if manufacturers price and/or negotiate such that the ICER for each new technology is as close as possible to AU$28,033 per QALY, then every additional AU$28,033 spent on a new technology will provide one QALY to the patients that benefit from the technology but will also forego one QALY among patients whose health care services are displaced; it follows that there is no *net* population benefit from adopting the new technology. If the policy objective is simply to ensure that new technologies do not displace more benefits than they provide, then the conventional supply-side approach to setting the threshold is consistent with this objective. However, if an alternative threshold exists at which net population benefits are positive, then the conventional supply-side approach is not consistent with an objective of maximizing total benefits across the population.

A further limitation of the conventional supply-side approach is that it considers only the benefits that arise to patients within the health care system. Under standard microeconomic theory, these benefits may be considered analogous to the 'consumer surplus' of the new technology. What is missing from conventional supply-side models is a consideration of 'producer surplus' - that is, profits to manufacturers arising from the adoption of new technologies.

The conventional demand-side approach also has important limitations. Not only does such an approach similarly ignore considerations of producer surplus, it also cannot be used to consider the consumer surplus arising from adopting new technologies within a budget constrained health care system; this is because a demand-side approach does not take into account any reductions in consumer surplus that result from the displacement of other health care services.

The need for an alternative approach

If policy makers wish to ensure that new technologies provide positive consumer surplus in the presence of strategic behaviour on the part of manufacturers, or if they are concerned about the allocation of both consumer and producer surplus, it follows that an alternative approach for determining the appropriate cost-effectiveness threshold is required.

The purpose of this paper is to propose a new conceptual model of the cost-effectiveness threshold. This proposal incorporates conventional supply-side and demand-side considerations into a single model, and builds upon existing approaches by incorporating strategic behaviour on the part of the manufacturers of new technologies. The proposed model allows decision makers to consider how different specifications of the threshold might impact upon the distribution of consumer and producer surplus arising from the adoption of new technologies. It also illuminates some potential additional avenues for future empirical research in this space.

Proposed model

The purpose of this section is to propose a new conceptual model of the threshold that accounts for two additional considerations:

1. Strategic pricing behaviour by the manufacturers of new technologies;

2. The impact of specifying a threshold upon ‘consumer’ and ‘producer’ surplus.

The proposed model is conceptual: it does not provide a ‘complete’ consideration of strategic behaviour or the determinants of consumer and producer surplus. Rather, the intention is to incorporate these considerations in a way that allows for consideration of potential departures from conventional approaches, while providing a framework to support future research.

After specifying some initial assumptions, the model will be developed first from the ‘consumer’ perspective (that of the health care system), and then from a ‘producer’ perspective (that of the manufacturers of new technologies supplied to the health care system). The models constructed under these two perspectives will then be combined. This will allow for consideration of the implications of specifying the threshold for the distribution of consumer and producer surplus. It will also allow for consideration of the ‘optimal threshold’ to use under each of a number of potential policy objectives.

Assumptions

In common with the conventional supply-side model, we will assume that:

1. There is a publicly funded health care system with a constrained budget;
2. There is an accepted measure of the ‘benefits’ that patients derive from health care;
3. Adopting new technologies displaces existing health care services, resulting in forgone ‘benefits’ for other patients.

As in existing papers, we will refer to the shadow price estimated in the conventional supply-side model as k (Claxton et al. 2011; Lomas et al. 2018). That is, k represents the relationship between marginal reductions in expenditure on existing health care services and forgone benefits for other patients.

For clarity, we will refer to the threshold to which the ICERs of new technologies are compared as λ. This allows us to distinguish this threshold from k; while the conventional supply-side model assumes that $\lambda = k$, in our model these may differ.

In addition, we will make the following assumptions:

4. The threshold is publicly stated by a decision maker and held constant over some time period, during which numerous new technologies are appraised;
5. The manufacturers of new technologies are strategic and ‘price to the threshold’, resulting in ICERs equal to the threshold;
6. Each manufacturer has a minimum ‘reserve price’ that must be met before supplying each new technology - this price is sufficient to cover the costs of production and an acceptable return on the manufacturer’s investment in research and development.

Each of assumptions 4-6 appears to be a reasonable approximation of real-world practice in many publicly funded health systems. For example, the UK’s National Institute for Health and Care Excellence (NICE) has publicly stated that its baseline threshold is £20,000 - £30,000 per QALY; this has remained constant for many years, during which numerous new technologies have been appraised. There is also empirical evidence of manufacturers ‘pricing to the threshold’ (although it may be a simplification to suppose that manufacturers price *exactly* to the threshold, as assumed here). In Canada, there are established processes for manufacturers and payers to negotiate on price following assessment of the cost-effectiveness evidence, allowing for manufacturers to set a high initial price and then negotiate the price down afterwards; the use of ‘risk-sharing’ schemes in the UK and elsewhere provides another mechanism by which a high initial ICER may effectively be negotiated down to the threshold. Furthermore, it is reasonable to expect a manufacturer to refuse to supply a technology if the price is too low to meet its production costs and achieve an acceptable return on its investment.

For any given ‘reserve price’ for a technology, there is an associated reserve ICER. This is because the price is a factor in determining the incremental cost of the technology, and in turn the ICER. Typically, a reduction in the price will lower the ICER (all other things equal), while an increase in the price will raise the ICER; it follows that if the technology will not be supplied below a specific ‘reserve price’, then, equivalently, it will not be supplied if the ICER is below a specific reserve ICER.

The precise relationship between the ‘reserve price’ and reserve ICER for a technology may be complicated, since (i) the ICER depends upon the incremental cost, which in turn depends on more than just the price of the technology, and also (ii) the price and other components of the incremental cost may be incurred across multiple years and be subject to discounting. However, for the purposes of this model, it is not necessary to understand this relationship in detail. Rather, we only need to make the following assumption:

8. There is a distribution of ‘reserve prices’ across new technologies, which gives rise to a distribution of reserve ICERs. Each new technology will be supplied to the health care system if, and only if, the specified threshold exceeds the reserve ICER.

Finally, we will assume that the distribution of reserve ICERs is broad, with some lying below k and others lying above k. This appears to be a reasonable assumption: violating this under a conventional supply-side approach (where $\lambda = k$) would result in *all* new technologies being adopted (if the distribution lies entirely below k) or *all* new technologies being rejected (if the distribution lies entirely above k), a clear departure from what is seen in practice. For simplicity, we will also not consider new technologies with zero or negative reserve ICERs, although very low reserve ICERs (close to zero) are permitted. Specifically, we will assume that:

9. ‘Reserve ICERs’ are continually distributed between zero and an ICER greater than k.

Consumer perspective

From the consumer perspective, the outcome of interest is assumed to be the ‘benefit’ provided by the adoption of new technologies. Since adopting a new technology provides direct ‘benefit’ to some patients, but also results in the displacement of other health services and so foregone ‘benefit’ for other patients, it is necessary to consider the impact on *net population ‘benefit’*.

Net population ‘benefit’ is considered to represent the ‘consumer surplus’ that arises from the adoption of new technologies. Note that no assumptions are made regarding the ‘units’ used to measure this ‘benefit’. In practice, the QALY is frequently used for this purpose, but this is not required for the implications of the proposed model to hold. For the remainder of this paper, we will refer to a generic measure of net population ‘benefit’, rather than any specific measure.

The consumer threshold curve

Figure 1 plots what we will hereafter refer to as the “consumer ‘threshold curve’” (or ‘CTC’). This ‘threshold curve’ represents the relationship between the threshold used to determine whether a technology is ‘cost-effective’ (λ), represented on the horizontal axis, and the net population ‘benefit’ (consumer surplus) derived from the adoption of new technologies, represented on the vertical axis.

Note that the curve plotted here represents a *stylized* CTC that satisfies the basic properties described below. Understanding the exact *shape* of the CTC in practice requires empirical research into the distribution of reserve ICERs across all new technologies; this, in turn, depends upon the health care system in question and is beyond the scope of this paper. The aim of this paper is simply to outline the *properties* that we would expect the CTC to have and some of the resulting implications.

Properties of the consumer’s threshold curve

The CTC bears some resemblance to the well-known ‘Laffer curve’, which describes the relationship between a tax rate and the resulting tax revenue (Fullerton 2016). The Laffer curve is anchored around two extreme points: a tax rate of 0% and a tax rate of 100%. At both of these anchor points the tax revenue is assumed to be zero. Between these points, tax revenue first increases and then decreases, such that there is some tax rate at which revenue is maximized. Although there is controversy around the *shape* of the Laffer curve and the point at which revenue is maximized in practice, the theoretical model nevertheless provides a useful insight: there is a tax rate above which revenues begin to fall, and so the optimal tax rate cannot be greater than this (Laffer 2004).

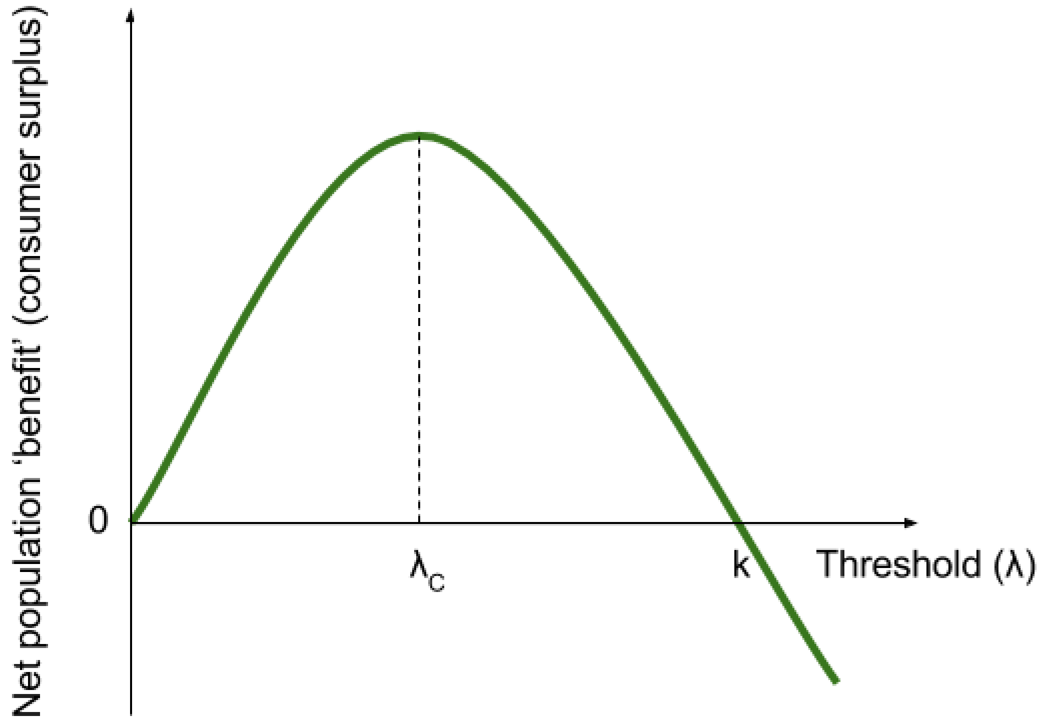


Figure 1: The 'consumer threshold curve', reflecting the relationship between the threshold (λ) and net population 'benefit' (consumer surplus)

The two anchor points of the consumer's 'threshold curve'

Similar to the Laffer curve, the CTC is anchored around two points on the horizontal axis: a threshold of zero and a threshold of k. Net population 'benefit' is zero at these anchor points. The reason for this is as follows:

- If the threshold is set equal to zero ($\lambda = 0$), no new technologies will be adopted. This is because the distribution of reserve ICERs lies entirely above zero (and hence λ), so the reserve ICER will not be met for any new technology. Since no new technologies will be adopted, it follows that no 'benefit' will be provided to patients, but also no 'benefit' will be displaced in other patients, such that the *net population 'benefit'* will be zero.

- If the threshold is set equal to k ($\lambda = k$), then some (but not all) new technologies will be adopted. This is because some new technologies have a reserve ICER below k, such that manufacturers will be prepared to supply these to the health care system, while other new technologies have a reserve ICER above k and so will not be supplied by manufacturers. For those new technologies that are supplied, some will have a lower reserve ICER than others; however, since all manufacturers strategically ‘price to the threshold’, the actual ICER for each supplied technology will equal k. Adopting these technologies will result in one unit of ‘benefit’ being displaced for every unit of ‘benefit’ provided, such that the *net population ‘benefit’* will be zero.

In common with the Laffer curve, the CTC begins from one anchor point, rises to a peak, and then falls to the second anchor point. However, unlike the Laffer curve, which is constrained between these two anchor points, the CTC intersects the horizontal axis at its second anchor point ($\lambda = k$) and extends beyond this point, becoming more negative with further increases in the threshold. The reasons for this particular shape are described below.

The shape between the anchor points

The shape of the CTC between the anchor points results from two countervailing effects that arise from changes in the threshold. These may be examined by considering a marginal increase in the threshold from λ_1 to λ_2, where both lie between zero and k ($0 < \lambda_1 < \lambda_2 < k$); this gives rise to the following effects:

1. The reserve ICER is now met for the subset of new technologies with reserve ICERs *between* λ_1 and λ_2. Previously these new technologies would not have been supplied by manufacturers, but following the marginal increase in the threshold these will now be provided. Since manufacturers strategically ‘price to the threshold’, each of these new technologies will be priced so that its ICER is equal to λ_2. Since $\lambda_2 < k$, the ‘benefit’ provided by each of these new technologies will exceed the ‘benefit’ forgone through the displacement of other health care services. This *additional supply* of new technologies therefore *increases* net population ‘benefit’.

2. Manufacturers of the subset of new technologies with reserve ICERs *below* λ_1, which would have been supplied even prior to the increase in the threshold, will strategically raise their prices until the ICER for each new technology equals λ_2. This increases the ‘benefit’ forgone through displacement, without providing any additional ‘benefit’ to patients. This *strategic pricing* therefore *decreases* net population ‘benefit’.

Between these anchor points, whether the CTC rises or falls following a marginal increase in the threshold depends upon the magnitude of each of these effects. If the first effect outweighs the second then the CTC will rise; if the second effect outweighs the first then the CTC will fall.

At the first anchor point ($\lambda = 0$), the first effect will be positive because a marginal increase in the threshold will cause a some new technologies to be supplied (those with very low reserve ICERs, such as some generics and other technologies with low marginal costs of production). The second effect will be zero because no new technologies have a reserve ICER below zero; since no new technologies are adopted when $\lambda = 0$, a marginal increase in the threshold from the first anchor point does not result in any strategic price increases. It follows that the first effect outweighs the second effect, causing the CTC to *rise* from the first anchor point.

With each successive marginal increase in the threshold, the positive impact of the first effect will tend to *diminish* because the additional new technologies supplied will be priced up to a progressively higher threshold, resulting in a relatively larger amount of forgone ‘benefit’ through displacement and hence a smaller increase in net population ‘benefit’. Meanwhile the negative impact of the second effect will tend to *grow* with increases in the threshold because the subset of new technologies with a reserve ICER below the threshold will also increase; with each successive marginal increase in the threshold, this growing subset of new technologies will be strategically priced up to the higher threshold, causing a greater amount of forgone ‘benefit’.

As a result of the diminishing impact of the first effect and the growing impact of the second effect, a threshold will be eventually be reached where the magnitude of these effects is equal. At this threshold, the gain in net population ‘benefit’ that arises from a marginal increase in the threshold (due to an increase in the supply of new technologies) is exactly offset by the loss in net population ‘benefit’ due to strategic pricing from manufacturers. Since the CTC is neither rising nor falling at this point, this represents the *peak* of the CTC. The threshold corresponding to this peak is hereafter referred to as the ‘optimal consumer threshold’ and denoted as λ_C. Net population ‘benefit’ (consumer surplus) is maximized at this threshold.

Further marginal increases in the threshold beyond λ_C result in a *reduction* in net population ‘benefit’, since the second effect now outweighs the first. Eventually, if the threshold is increased all the way to the second anchor point ($\lambda = k$), net population ‘benefit’ will reduce to zero and the CTC will intersect the horizontal axis. Net population ‘benefit’ is zero at the second anchor point because manufacturers will price new technologies so that each ICER equals k; as noted earlier, this will result in every unit of ‘benefit’ produced by new technologies being exactly offset by a unit of ‘benefit’ forgone by other patients due to displacement.

The shape beyond the second anchor point

When the threshold is increased beyond the second anchor point ($\lambda > k$), the first of the two countervailing effects described above begins to impact upon net population ‘benefit’ in the *opposite* direction, such that *both* effects now act to *diminish* net population ‘benefit’.

This is because each additional new technology supplied following a marginal increase in the threshold will be strategically priced to have an ICER above k, such that adoption will cause more ‘benefit’ to be forgone through displacement than will be gained by patients. As before, a marginal increase in the threshold will also cause the manufacturers of new technologies that would have been supplied at the previous threshold to price up to the higher threshold, causing additional displacement. It follows that increases in the threshold beyond k will unambiguously cause net population ‘benefit’ to *fall*.

Further considerations

The specific curve plotted in **Figure 1** is just one of a set of possible curves that satisfy the properties above. In practice, the shape of the CTC might differ from that in **Figure 1** for one or more reasons, including (but not limited to) the following:

1. The *skewness* of the distribution of reserve ICERs for new technologies. If a greater proportion of new technologies have very low reserve ICERs, then we might expect the peak of the CTC to be shifted to the left, resulting in a lower optimal consumer threshold. Conversely, if a greater proportion of new technologies have very high reserve ICERs, then the peak might be shifted to the right, resulting in a higher optimal consumer threshold. Nevertheless, regardless of the distribution of reserve ICERs, the properties of the CTC require that the optimal consumer threshold lies between zero and k.

2. The *density* of the distribution of reserve ICERs for new technologies. For any given threshold, the greater the density of the distribution of reserve ICERs below this threshold, the greater the number of new technologies adopted and the greater the *magnitude* of the gain or loss in net population benefit. This will cause a vertical stretch of the CTC in the vertical plane, but will not impact upon the location of the peak along the horizontal axis, and hence will not affect the optimal consumer threshold.

3. The magnitude of the *'benefit'* provided by adopted new technologies. Among new technologies with the same ICER, some may provide greater 'benefit', at a correspondingly higher price, than other new technologies. Adopting new technologies which provide greater 'benefit' will have a greater impact on net population 'benefit' than adopting new technologies with identical ICERs but which provide lower 'benefit'.

Producer perspective

From the producer perspective, the outcome of interest is the profit provided to manufacturers by the adoption of new technologies. For the purposes of this model, it will be assumed that 'producer surplus' reflects the profits that arise to manufacturers that supply new technologies to the health care system.[1]

It should be noted that this definition of 'producer surplus' does not consider losses incurred by manufacturers of new technologies which are *not* supplied to the health care system. Since this is a potentially controversial assumption, a justification is provided below. A modified model which considers the implications of modifying this assumption is provided in the Appendix.

Justification for excluding manufacturers who do *not* supply new technologies to the health care system

Supplying the health care system with new technologies is a competitive process. When a cost-effectiveness threshold is used to determine which new technologies are adopted and which are not, it is inevitable that some manufacturers will lose out. Where manufacturers have invested in research and development of new technologies, they may incur losses as a result.

When considering the impact of the threshold upon consumer and producer surplus, there are good reasons why an agency may wish to consider, when calculating producer surplus, *only* the profits arising to manufacturers who supply the health care system with new technologies, and not the losses incurred by *non-supplying* manufacturers. Considering these losses would serve to lower the overall producer surplus. In cases where the agency desires that a specific share of the overall surplus be allocated to producers, or adopts a constraint that producer surplus cannot be negative, it follows that considering losses incurred by *non-supplying* manufacturers may require an *increase* in the threshold to satisfy such an objective. This, in turn, will cause a reduction in net population 'benefit' (consumer surplus) if the threshold is raised above λ_C.

It is questionable whether the agency responsible for specifying the cost-effectiveness threshold has any obligation to diminish net population 'benefit' (consumer surplus) in order to support manufacturers who have failed to develop new technologies that provide additional 'benefit' to the health care system. It may instead be considered preferable to foster competition between manufacturers to supply new technologies at ICERs below the desired threshold. Those manufacturers who invest productively and manufacture new technologies efficiently will tend to have lower reserve ICERs than those manufacturers who are wasteful in their research and

[1] It is assumed here, for simplification, that each new technology is supplied by a different 'manufacturer'. In practice, a single manufacturer may supply multiple new technologies to the health care system, and may develop several other new technologies that are *not* supplied to the health care system. In this case, the 'producer surplus' considered in this model would reflect the profits associated with *only* those new technologies that each manufacturer supplies to the health care system.

development, adopt inefficient manufacturing techniques, or require unrealistically high rates of return before supplying their technologies to the health care system. Under the model proposed in this paper, the most efficient manufacturers - those who develop new technologies with the lowest reserve ICERs - receive *super-normal* profits when they supply their new technologies to the health care system, since they can strategically price up to the threshold. This provides a clear incentive for manufacturers to improve efficiency and develop new technologies that provide as large a ‘benefit’ to the health care system as possible at as low a cost as possible, both factors that will contribute to lowering the reserve ICER. Those manufacturers who are slightly less efficient, and develop new technologies with slightly higher reserve ICERs, will receive *smaller* but still *super-normal* profits when they strategically price to the threshold. Manufacturers with reserve ICERs equal to the threshold will make no super-normal profits, but also no losses. Losses will be incurred only by those manufacturers who are the least efficient at producing ‘benefit’ for the health care system, since their technologies will have reserve ICERs above the specified threshold and so will not be adopted.

Incorporating the losses incurred by *non-supplying* manufacturers into the consideration of producer surplus, resulting in a higher threshold, would diminish the incentives described above. Manufacturers who fail to develop new technologies that provide additional ‘benefit’ to the health care system may nevertheless be rewarded if the threshold is increased as a result of considering these losses. In addition, since manufacturers will now price up to a higher threshold, much of the gain in producer surplus will be enjoyed by manufacturers who would have supplied their technologies under the existing threshold; increasing the threshold serves to increase the already super-normal profits enjoyed by these manufacturers. Furthermore, manufacturers with reserve ICERs above the increased threshold will still incur the same losses as before.

It follows that raising the threshold is not necessarily an effective or efficient means of mitigating the losses incurred by manufacturers whose new technologies are *not* adopted by the health care system, are diminishes the incentives for manufacturers to efficiently develop new technologies that provide ‘benefit’ to the health care system. Limiting the consideration of producer surplus to *only* the profits enjoyed by manufacturers who *supply* new technologies to the health care system avoids these issues and encourages greater competition among manufacturers in the market to provide new technologies to the health care system.

The producer threshold curve

Figure 2 plots what we will hereafter refer to as the "producer 'threshold curve'" (or 'PTC'). This 'threshold curve' represents the relationship between λ, represented on the horizontal axis, and 'manufacturer profit' (producer surplus) arising from the supply of new technologies, represented on the vertical axis.

As with the CTC in **Figure 1**, the curve plotted in **Figure 2** represents a *stylized* PTC that satisfies the basic properties described below. Understanding the exact *shape* of the PTC in practice requires empirical research into the distribution of reserve ICERs across all new technologies. The aim of this paper is simply to outline the *properties* that we would expect the PTC to have and some of the resulting implications.

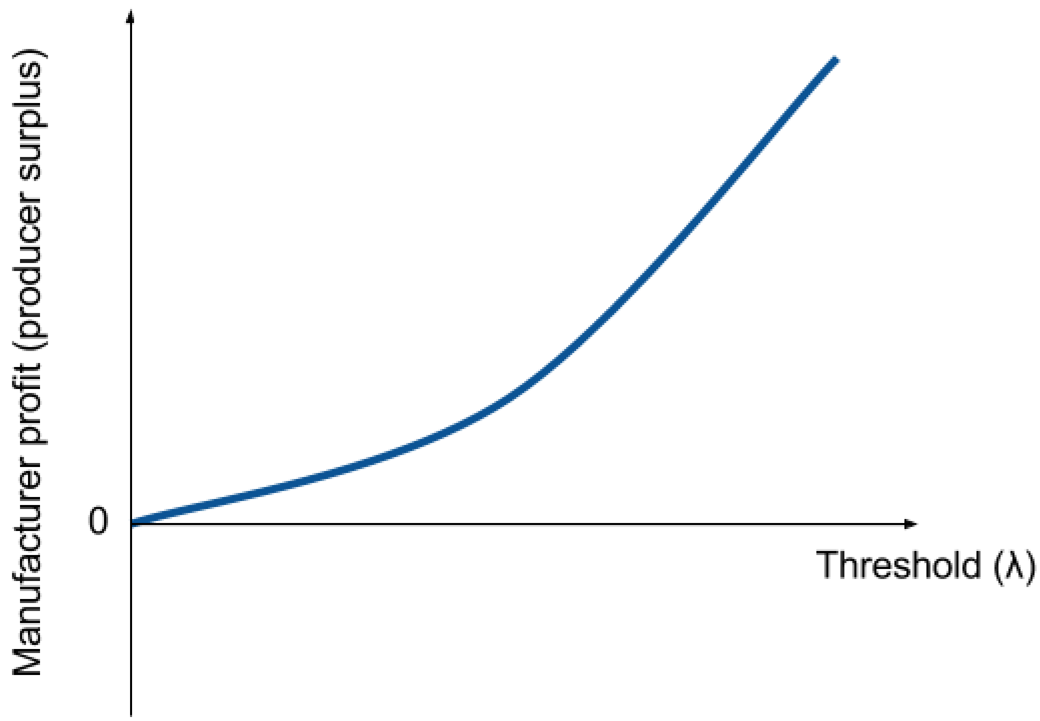


Figure 2: The 'producer threshold curve', reflecting the relationship between the threshold (λ) and manufacturer profit (producer surplus)

Properties of the producer's 'threshold curve'

The PTC has the following properties:

- If the threshold is set equal to zero ($\lambda = 0$), manufacturer profit is also *zero*. This is because no new technologies are adopted by the health care system, and hence no manufacturers profit from supplying new technologies to the health care system.

- As the threshold increases above zero, manufacturer profit becomes *positive*. This is because new technologies with reserve ICERs below the threshold are now supplied to the health care system. Manufacturers of these new technologies strategically price up the threshold, resulting in *super-normal* profits.

- With further increases in the threshold, manufacturer profit will *unambiguously and continuously increase*. This is due to two effects, both of which cause profit to increase with the threshold. First, the reserve ICERs will be met for additional new technologies, causing them to be supplied to the health care system; each is strategically priced up to the threshold, resulting in super-normal profits for their manufacturers. Second, all new technologies with lower reserve ICERs (those that would be supplied even without an increase in the threshold) will now be strategically priced up to the higher threshold, resulting in additional profit for manufacturers.

It follows that the PTC lies entirely above the horizontal axis and continues to increase (without limit) with increases in the threshold. Even if the threshold is already so high that the reserve ICER is met for *all* new technologies, additional increases in the threshold will increase profits by allowing all manufacturers to price up to a higher threshold.

Combining the consumer and producer perspectives

We will now combine the 'consumer' and 'producer' perspectives that were considered over previous sections. This will allow for consideration of the distribution of 'consumer surplus' and 'producer surplus' that arises under different specifications of the threshold. This, in turn, will allow for consideration of the 'optimal threshold' to adopt under different policy objectives.

Converting 'benefit' and 'profit' into a common metric

Unless the measure of 'benefit' from the consumer perspective is already specified in monetary terms, a requirement for aggregating consumer and producer surplus is to convert each into a common metric. Whether this is done by converting consumer surplus into monetary terms or by converting producer surplus into units of consumer 'benefit' is immaterial; the most important and challenging task is identifying the relevant *conversion rate*.

A conventional *demand-side* approach to determining the threshold provides a natural source for such a conversion rate. Demand-side approaches typically involve estimation of the *monetary value* of a unit of 'benefit'; such estimates may be used directly to convert net population 'benefit' into monetary terms, or to convert manufacturer profit into units of 'benefit', allowing both to be considered in a common metric.

Because there are competing methodologies for empirically estimating demand-side thresholds, and since any estimate is context dependent, for the purposes of this paper we will not assume any particular conversion rate between consumer and producer surplus. We will therefore constrain our consideration of the implications of our model to those which arise regardless of the conversion rate used.

Comparing the consumer and producer threshold curves

Figure 3 plots the CTC and PTC from **Figure 1** and **Figure 2** respectively on a single graph. This figure reveals the inherent tension between consumer and producer interests in any specification of the threshold. Setting the threshold equal to the 'optimal consumer threshold' (λ_C) maximizes consumer surplus, with a net population 'benefit' of C*. At this threshold, producer surplus is positive, with a manufacturer profit of P_C, but is not maximized; producer surplus can be expanded by increasing the threshold, but this comes at the expense of diminished consumer surplus. If the threshold is increased to k then producer surplus rises to P_k but consumer surplus falls to zero. At higher thresholds, producer surplus increases further but consumer surplus becomes negative.

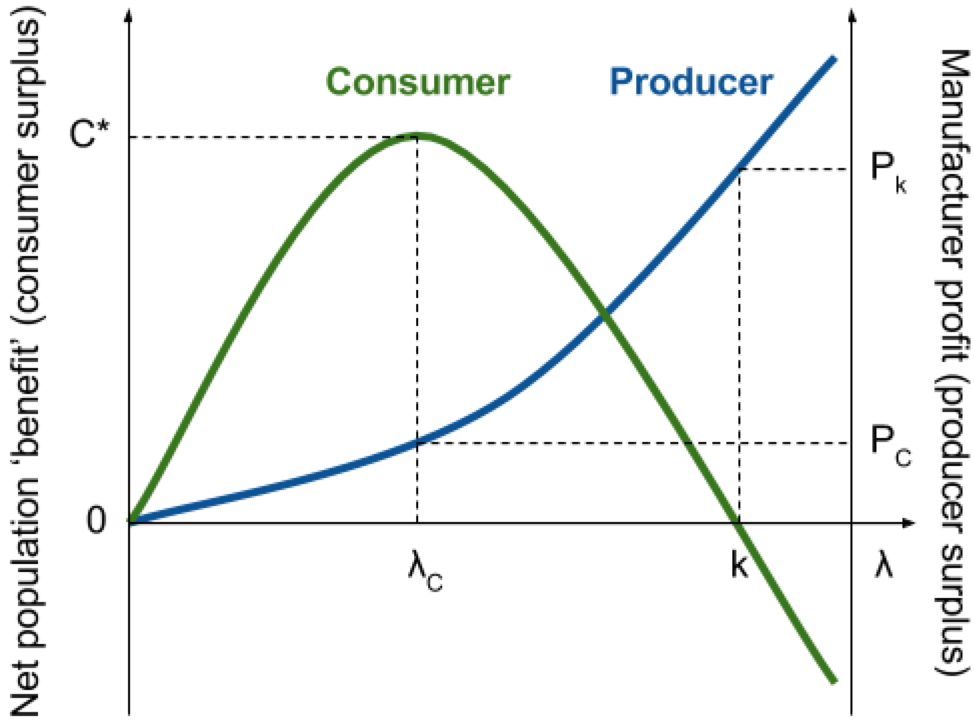


Figure 3: Consumer and producer threshold curves, reflecting the relationship between the threshold (λ), net population 'benefit' (consumer surplus) and manufacturer profit (producer surplus)

The relationship between changes in the threshold and the sign and direction of change for consumer surplus, producer surplus, and the *combined surplus* are summarized in **Table 1**.

Below a threshold of λ_C, consumer and producer surplus are both positive and both increase with the threshold; it follows that the combined surplus is also positive and increasing, regardless of the conversion rate used. An implication of this is that there is no reason to specify a threshold below λ_C, since a higher threshold would benefit both consumers and producers. (This implication is very similar to that from the Laffer curve, where the area to right of the peak is considered 'off-limits' since an identical tax revenue may be collected with a lower tax rate; here, the area to the left of the peak of the CTC may be considered 'off-limits' since consumer and producer surplus can both be increased by specifying a higher threshold).

Above a threshold of λ_C, but below a threshold of k, consumer surplus decreases but remains positive, while producer surplus increases still further. It follows that the combined surplus is also positive across this range, although its direction of change is ambiguous since this depends upon the *conversion rate* between consumer and producer surplus.

Above a threshold of k, consumer surplus becomes negative and continues decreasing thereafter, while producer surplus continues increasing. It follows that both the sign and the direction of change of the combined surplus are ambiguous across this range, since both now depend upon the conversion rate between consumer and producer surplus.

Threshold range	**Property**	**Consumer surplus**	**Producer surplus**	**Combined surplus**
$0 < \lambda < \lambda_C$	Sign	Positive	Positive	Positive
	Direction of change	Increasing	Increasing	Increasing
$\lambda_C < \lambda < k$	Sign	Positive	Positive	Positive
	Direction of change	Decreasing	Increasing	Ambiguous
$\lambda > k$	Sign	Negative	Positive	Ambiguous
	Direction of change	Decreasing	Increasing	Ambiguous

Table 1: Sign and direction of change for consumer surplus, producer surplus and combined surplus across different threshold ranges

Specification of 'optimal' thresholds

Given the inherent tension between consumer and producer interests, the specification of an 'optimal' threshold is a difficult task. It depends crucially upon the policy objective adopted, specifically the desired distribution of the combined surplus between consumers and producers.

The determination of this objective is a matter for policy makers, and so no specific objective will be assumed here. Nevertheless, it is useful to explore the implications of alternative objectives for the specification of the optimal threshold. We will therefore consider seven possible policy objectives and the optimal thresholds associated with each. In cases where we are unable to precisely determine the optimal threshold, we will instead specify a range in which the optimal threshold must lie, given the properties of each threshold curve described earlier.

Policy objectives

A number of possible policy objectives exist, including (but not limited to) the following:

1. Maximize *consumer* surplus;
2. Maximize *consumer* surplus, subject to consumer and producer surplus each being *non-negative*;
3. Maximize *consumer* surplus, subject to *producer* surplus comprising a *guaranteed proportion of the combined surplus* and also subject to each being *non-negative*;
4. Maximize *producer* surplus;
5. Maximize *producer* surplus, subject to consumer and producer surplus each being *non-negative*;
6. Maximize *producer* surplus, subject to *consumer* surplus comprising a *guaranteed proportion of the combined surplus* and also subject to each being *non-negative*;
7. Maximize the *combined* surplus (consumer and producer surplus);
8. Maximize the *combined* surplus, subject to each being *non-negative*.

Satisfying objectives 1, 4 or 7 may require consumer or producer surplus to be *negative*, an outcome that might not be considered reasonable by patients or manufacturers. For example, it might be considered unreasonable to expect the health care system to adopt technologies that diminish net population 'benefit' in order to increase manufacturer profit, or to expect manufacturers to supply new technologies at a loss in order to provide 'benefit' to patients.

Objectives 2, 3, 5, 6 and 8 address these concerns by requiring that both consumer and producer surplus be non-negative. Objectives 3 and 6 also incorporate a concern for the *proportion* of the combined surplus that is allocated to consumers or producers. Note that the maximum surplus achievable may be less in the presence of each of these constraints.

Policy objective	Optimal threshold, or range containing optimal threshold	Comments
'Maximize *consumer* surplus'	$\lambda^* = \lambda_C$	Consumer surplus is maximized by specifying a threshold of λ_C.
'Maximize *consumer* surplus, subject to consumer and producer surplus each being *non-negative*'	$\lambda^* = \lambda_C$	At a threshold of λ_C, consumer surplus is maximized and producer surplus is positive.
'Maximize *consumer* surplus, subject to *producer* surplus comprising a *guaranteed proportion of the combined surplus* and also subject to each being *non-negative*'	$\lambda_C \le \lambda^* \le k$	The proportion of the combined surplus allocated to producers increases above λ_C. If producer surplus comprises the required proportion of the combined surplus at λ_C then this is the optimal threshold. If not, the threshold should be progressively increased until the required proportion is achieved
'Maximize *producer* surplus'	$\lambda^* = \infty$	Producer surplus is maximized with an infinitely high threshold.
'Maximize *producer* surplus, subject to consumer and producer surplus each being *non-negative*'	$\lambda^* = k$	Since producer surplus increases with the threshold, and consumer surplus is negative at any threshold above k, this objective is satisfied by specifying a threshold of k.
'Maximize *producer* surplus, subject to *consumer* surplus comprising a *guaranteed proportion of the combined surplus* and also subject to each being *non-negative*'	$0 < \lambda^* \le k$ OR *N/A*	As above, except that the optimal threshold is derived by progressively lowering the threshold from k until the required proportion of consumer surplus is achieved. If the threshold is lowered to zero and this proportion is still not achieved then no threshold exists that satisfies this objective.
'Maximize the *combined* surplus (consumer and producer surplus)'	$\lambda^* > \lambda_C$	Consumer and producer surplus both increase with the threshold up to λ_C. Above λ_C, consumer surplus falls and producer surplus increases. The optimal threshold therefore depends upon the shape of each threshold curve but must exceed λ_C.
'Maximize the *combined* surplus, subject to each being *non-negative*'	$\lambda_C < \lambda^* \le k$	Since consumer and producer surplus both increase with the threshold up to λ_C, but consumer surplus is negative above k, the optimal threshold must lie between λ_C and k.

Table 2: Optimal threshold (λ*), or range containing optimal threshold, for each objective

‘Optimal’ thresholds for each objective

Table 2 reports the optimal threshold (λ^*), or the range containing the optimal threshold, for each of the eight policy objectives considered above.

Where the objective is to maximize *consumer* surplus, the optimal threshold is λ_C. This finding holds if there is also a desire for producer surplus to be non-negative, since producer surplus is positive at this threshold. If, in addition, there is a desire that producer surplus comprise a guaranteed proportion of the combined surplus, this may require that the optimal threshold be increased above λ_C (but no higher than k) until this proportion is reached.

Where the objective is to maximize *producer* surplus, the optimal threshold is infinitely high. However, this results in negative consumer surplus; if there is also a desire for consumer surplus to be non-negative, then the optimal threshold is k. If, in addition, there is a desire that consumer surplus comprise a guaranteed proportion of the combined surplus, this may require that the optimal threshold be lowered below k until this proportion is reached.

Finally, where the objective is to maximize the *combined* surplus, the optimal threshold lies somewhere above λ_C, with its *precise* location dependent upon the shape of each threshold curve and the conversion rate between consumer and producer surplus. If there is also a desire that both consumer and producer surplus be non-negative, the optimal threshold lies somewhere above λ_C but no higher than k.

Discussion

This paper has proposed a new conceptual model of the cost-effectiveness threshold that accounts for strategic behaviour on the part of manufacturers and allows for consideration of 'optimal' thresholds under various policy objectives regarding the distribution of consumer and producer surplus. This proposal combines elements of conventional supply-side and demand-side approaches into a single model: the conventional supply-side threshold (k) forms one of two anchors of the consumer threshold curve, while the conventional demand-side threshold is used to convert consumer and producer surplus into a common metric.

Despite building upon these familiar foundations, the integration of strategic pricing behaviour into the model has resulted in implications that depart from those of conventional supply-side and demand-side approaches. The conventional supply-side approach has been shown to be consistent with only one of the eight policy objectives considered here: that of maximizing *producer* surplus subject to the constraint that consumer surplus is non-negative. This objective differs substantially from that which the supply-side approach is widely assumed to satisfy: maximizing net population 'benefit' (i.e., *consumer* surplus). It is debatable whether proponents of a conventional supply-side approach to the threshold would be in favour of adopting such an objective over some of the other objectives considered here; if an alternative policy objective is adopted, k is not generally the optimal threshold to specify.

If the policy objective is to maximize *consumer* surplus, then a *lower* threshold than k should be used: we define this as the 'optimal consumer threshold', λ_C. Specifying λ_C requires an understanding of the shape of the consumer threshold curve; this in turn requires an empirical estimate of k and also an understanding of the distribution of reserve ICERs across all new technologies. Although a number of recent studies have published empirical estimates of k, the latter consideration has not been subject to any empirical research to date. There is, therefore, a need for broadened empirical research if maximizing consumer surplus is the policy objective.

If policy is instead focussed upon maximizing *producer* surplus, then there is no limit as to how high the threshold should be set. Alternatively, if the focus is on maximizing producer surplus subject to a constraint that consumer surplus be non-negative, then the optimal threshold is k, since this is the highest threshold at which consumer surplus is non-negative.

Limitations

The model proposed here is conceptual and makes a number of strong assumptions. It is expected that some of these assumptions may be explored and relaxed in future research. In particular, we have assumed that manufacturers are *perfectly* strategic, always increasing prices up to the threshold and never behaving in a way that does not maximize profits. This might not be entirely accurate.

Possible future advancements to the model include relaxing the assumption that the policy maker sets only a single threshold, allowing for future changes in the price of technologies (as drugs lose patent protection and generic competitors enter the market), and more complex reimbursement mechanisms (such as risk sharing schemes) that may also be subject to strategic behaviour on the part of manufacturers.

Policy implications

The proposed model suggests that the optimal threshold is conditional upon a number of factors, including the policy objective, the conversion rate between consumer and producer surplus (a demand-side consideration), the shadow price of the health care budget constraint (a supply-side consideration), the distribution of reserve ICERs, and other factors that may affect the consumer and producer threshold curves.

Based on recent empirical estimates of supply-side thresholds (£12,936, €24,870 and AU$28,033 per QALY in England, Spain and Australia, respectively), the proposed model implies that, if decision makers in these countries have a primary concern for maximizing consumer surplus, then thresholds *lower* than these should be specified in practice. The use of *higher* thresholds is consistent with an objective of maximizing *producer* surplus, subject to a weak concern for consumer surplus that serves only to limit the extent to which it is negative.

References

Claxton, K. et al., 2011. Discounting and decision making in the economic evaluation of health-care technologies. *Health economics*, 20(1). Available at: http://dx.doi.org/10.1002/hec.1612.

Claxton, K. et al., 2015. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health technology assessment*, 19(14). Available at: http://dx.doi.org/10.3310/hta19140.

Eckermann, S. & Pekarsky, B., 2014. Can the real opportunity cost stand up: displaced services, the straw man outside the room. *PharmacoEconomics*, 32(4), pp.319–325. Available at: http://dx.doi.org/10.1007/s40273-014-0140-3.

Edney, L.C. et al., 2017. Estimating the Reference Incremental Cost-Effectiveness Ratio for the Australian Health System. *PharmacoEconomics*. Available at: http://dx.doi.org/10.1007/s40273-017-0585-2.

Fullerton, D., 2016. Laffer curve. In *The New Palgrave Dictionary of Economics*. pp. 839–841. Available at: http://dx.doi.org/10.1057/9780230226203.0922.

Laffer, A.B., 2004. The Laffer curve: Past, present, and future. *Heritage Foundation Backgrounder*, (1765), pp.1176–1196. Available at: http://walkerd.people.cofc.edu/400/Sobel/2A-7.%20Laffer%20-%20The%20Laffer%20Curve.pdf.

Lomas, J. et al., 2018. Resolving the “Cost-Effective but Unaffordable” Paradox: Estimating the Health Opportunity Costs of Nonmarginal Budget Impacts. *Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research*. Available at: http://www.sciencedirect.com/science/article/pii/S1098301517336136.

Olsen, J.A., 2017. Costs and the cost-effectiveness threshold. In *Principles in Health Economics and Policy*. Oxford: Oxford University Press. Available at: http://www.oxfordscholarship.com/10.1093/oso/9780198794837.001.0001/oso-9780198794837-chapter-19.

Paulden, M., McCabe, C. & Karnon, J., 2014. Achieving allocative efficiency in healthcare: nice in theory, not so NICE in Practice? *PharmacoEconomics*, 32(4), pp.315–318. Available at: http://dx.doi.org/10.1007/s40273-014-0146-x.

Paulden, M., O’Mahony, J. & McCabe, C., 2017. Determinants of Change in the Cost-effectiveness Threshold. *Medical decision making: an international journal of the Society for Medical Decision Making*, 37(2), pp.264–276. Available at: http://dx.doi.org/10.1177/0272989X16662242.

Pekarsky, B., 2012. Trust, constraints and the counterfactual: Reframing the political economy of new drugs. *University of Adelaide*. Available at: http://link.springer.com/article/10.1007/978-3-319-08903-4_3.

Remme, M., Martinez-Alvarez, M. & Vassall, A., 2017. Cost-Effectiveness Thresholds in Global Health: Taking a Multisectoral Perspective. *Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research*, 20(4), pp.699–704. Available at: http://dx.doi.org/10.1016/j.jval.2016.11.009.

Revill, P. et al., 2015. *Cost-effectiveness thresholds: guiding health care spending for population health improvement*, Centre for Health Economics, University of York. Available at: http://www.idsihealth.org/wp-content/uploads/2015/01/CE-Thresholds-iDSI-Working-Group-Final-Report.pdf.

Vallejo-Torres, L. et al., 2016. On the Estimation of the Cost-Effectiveness Threshold: Why, What, How? *Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research*, 19(5), pp.558–566. Available at: http://dx.doi.org/10.1016/j.jval.2016.02.020.

Vallejo-Torres, L., García-Lorenzo, B. & Serrano-Aguilar, P., 2017. Estimating a cost-effectiveness threshold for the Spanish NHS. *Health economics*. Available at: http://dx.doi.org/10.1002/hec.3633.

Appendix

Considering losses for non-supplying manufacturers

If the agency that specifies the cost-effectiveness threshold wishes to consider, within the calculation of producer surplus, the losses incurred by manufacturers that do *not* supply new technologies to the health care system, then this will have implications for the specification of the 'optimal' threshold under each policy objective. The purpose of this Appendix is to explore how the 'optimal' threshold might change when these losses are taken into consideration.

The producer threshold curve

When these losses are taken into consideration, the producer threshold curve (PTC) is no longer entirely above the horizontal axis. This is because, at very low thresholds, these losses exceed the small profits enjoyed by manufacturers who supply new technologies to the health care system; it follows that the PTC is negative at very low thresholds.

As the threshold increases, profits for manufacturers who supply new technologies to the health care system increase, and eventually exceed the losses incurred by other manufacturers. At this point, the PTC intersects the horizontal axis and the producer surplus becomes positive. We will refer to the threshold corresponding to this point as the 'minimum producer threshold' (λ_P), since this is the minimum threshold at which producer surplus is non-negative.

The implication of considering these losses for the 'optimal' threshold depends upon *where* the PTC intersects the horizontal axis, and hence the value of λ_P relative to λ_C and k. There are three possible scenarios:

1. The PTC intersects the horizontal axis to the *left* of the peak of the CTC; it follows that the minimum producer threshold is less than the optimal consumer threshold, which in turn is less than k ($\lambda_P < \lambda_C < k$). This implies that these losses are relatively *small*.
2. The PTC intersects the horizontal axis to the *right* of the peak of the CTC, but to the *left* of where the CTC intersects the horizontal axis; it follows that the minimum producer threshold exceeds the optimal consumer threshold but is still less than k ($\lambda_C < \lambda_P < k$). This implies that these losses are *moderate*.
3. The PTC intersects the horizontal axis to the *right* of where the consumer's 'threshold curve' intersects the horizontal axis; it follows that the minimum producer threshold is higher than k, which in turn exceeds the optimal consumer threshold ($\lambda_C < k < \lambda_P$). This implies that these losses are relatively *large*.

The 'optimal' thresholds under each scenario, for each policy objective, are summarized in

Table 3. The modified PTCs are plotted in **Figure 4**, **Figure 5** and **Figure 6**, respectively.

Policy objective	**Optimal threshold, or range containing optimal threshold**		
	Scenario 1	**Scenario 2**	**Scenario 3**
'Maximize *consumer* surplus'	$\lambda^* = \lambda_C$	$\lambda^* = \lambda_C$	$\lambda^* = \lambda_C$
'Maximize *consumer* surplus, subject to consumer and producer surplus each being *non-negative*'	$\lambda^* = \lambda_C$	$\lambda^* = \lambda_P$	*N/A*
'Maximize *consumer* surplus, subject to *producer* surplus comprising a *guaranteed proportion of the combined surplus* and also subject to each being *non-negative*'	$\lambda_C \le \lambda^* \le k$	$\lambda_P \le \lambda^* \le k$	*N/A*
'Maximize *producer* surplus'	$\lambda^* = \infty$	$\lambda^* = \infty$	$\lambda^* = \infty$
'Maximize *producer* surplus, subject to consumer and producer surplus each being *non-negative*'	$\lambda^* = k$	$\lambda^* = k$	*N/A*
'Maximize *producer* surplus, subject to *consumer* surplus comprising a *guaranteed proportion of the combined surplus* and also subject to each being *non-negative*'	$\lambda_P < \lambda^* \le k$	$\lambda_P < \lambda^* \le k$	*N/A*
'Maximize the *combined* surplus (consumer and producer surplus)'	$\lambda^* > \lambda_C$	$\lambda^* > \lambda_C$	$\lambda^* > \lambda_C$
'Maximize the *combined* surplus, subject to each being *non-negative*'	$\lambda_C < \lambda^* \le k$	$\lambda_P < \lambda^* \le k$	*N/A*

Table 3: Optimal threshold (λ*), or range containing optimal threshold, for each objective when losses for non-supplying manufacturers are considered

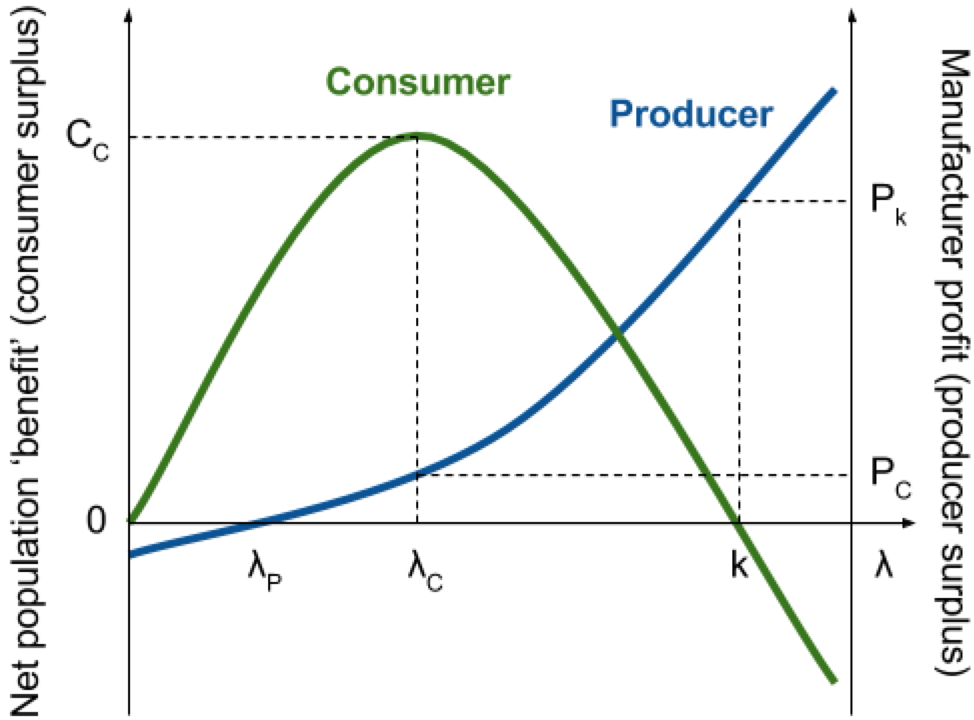


Figure 4: Consumer and producer threshold curves when losses for non-supplying manufacturers are considered and these losses are small ($\lambda_P < \lambda_C < k$) (Scenario 1)

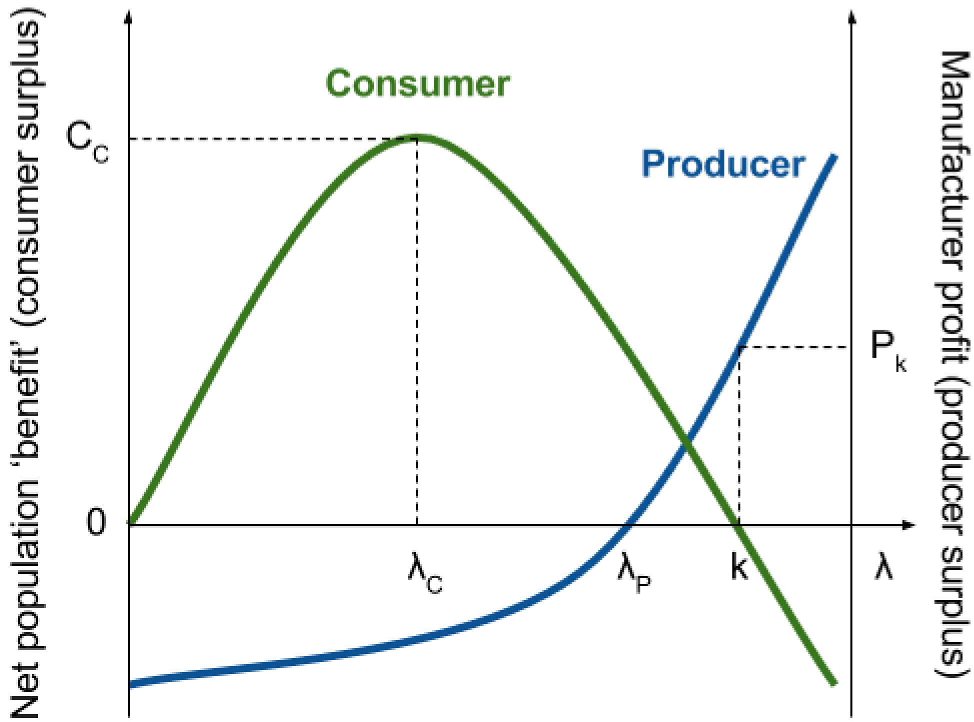


Figure 5: Consumer and producer threshold curves when losses for non-supplying manufacturers are considered and these losses are moderate ($\lambda_C < \lambda_P < k$) (Scenario 2)

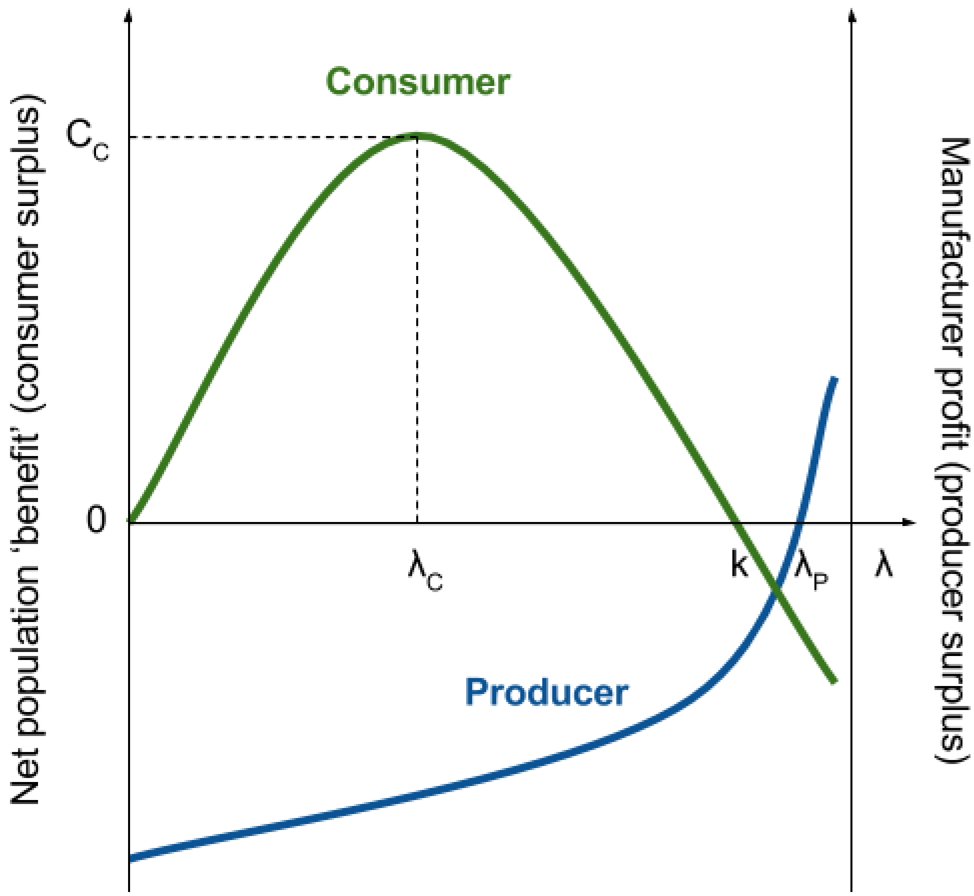


Figure 6: Consumer and producer threshold curves when losses for non-supplying manufacturers are considered and these losses are large (λ_C < k < λ_P) (Scenario 3)