Cost-effectiveness thresholds: guiding health care spending for population health improvement

Paul Revill, Jessica Ochalek,
James Lomas, Ryota Nakamura,
Beth Woods, Alex Rollinger, Marc Suhrcke,
Mark Sculpher, Karl Claxton.

November 2015

**A report by the Centre for Health Economics,
University of York,
for the International Decision Support Initiative (iDSI)**

iDSI Methods Working Group report

This report was produced as part of the Research arm of the International Decisions Support Initiative (iDSI); in a collaboration between the University of Glasgow, the London School of Hygiene and Tropical Medicine, Erasmus University, Rotterdam, and the Centre for Health Economics, University of York.

In the first phase of iDSI (2014-2016), Methods Working Groups were established to investigate three areas of economic evaluation for which additional methods guidance for analysts and policymakers in low- and middle-income countries was deemed particularly valuable. These linked directly to specific Principles for the practice of economic evaluation detailed in the iDSI Reference Case. The three research areas were identified by policymakers in low- and middle-income countries as being particularly challenging and requiring additional methods guidance in order to support the realization of their corresponding Reference Case Principles.

This report details the findings from the investigation into cost-effectiveness thresholds for economic evaluation - linked to Principle 10: Impact on Other Constraints and Budget Impact – "*The impact of implementing the intervention on the health budget and on other constraints should be identified clearly and separately*".

Additional information about the findings from the Methods Working Group can be viewed at: www.idsihealth.org/knowledge_base.

Acknowledgements

This report is financially supported by the Bill and Melinda Gates Foundation, the UK Department for International Development, and the Rockerfeller Foundation, through the International Decision Support Initiative (iDSI).

Lay summary

Health care systems need to be designed in ways that support the generation and pooling of health care resources and the allocation of those resources to deliver health benefits to constituent populations.

Cost-effectiveness analysis (CEA) is a set of tools that offer answers to the question: 'How should available resources be allocated to maximize population health benefits'. Health benefits can be defined in various ways but in most health care systems a principal objective is improvement in population health, as measured using metrics that combine length and quality of life such as quality adjusted life years (QALYs) gained or disability adjusted life years (DALYs) averted.

CEA seeks to identify which drugs, healthcare technologies, programmes or other interventions offer greater health benefits when funded than health benefits forgone as resources are not then available to fund other priorities. In other words, CEA is about identifying those interventions that offer health benefits greater than their opportunity costs.

Applied CEA studies need to identify (i) the health benefits offered by any intervention being evaluated; (ii) the additional costs imposed on a limited healthcare budget; and (iii) the opportunity costs (i.e. health benefits forgone) due to a commitment of resources to an intervention's provision. An intervention can only reasonably be deemed "cost-effective" if its benefits outweigh the opportunity costs of health benefits forgone.

The opportunity costs in terms of forgone health benefits are reflected in most CEAs by using a cost-effectiveness threshold (CET). The CET is a measure of the 'cost per unit of health benefit (e.g. cost per QALY gained/DALY-averted) forgone'. The forgone health is the health associated with the interventions that would be displaced, were the intervention under evaluation to be funded. For example if the CET is $1,000 per QALY then for every $1,000 spent on an intervention, this $1,000 can no longer be spent on other health care priorities resulting in a reduction in population health of 1 QALY.

In most health care systems the appropriate CET is not readily apparent but depends upon the particular funding arrangements in the system and the health benefits of other interventions with claims on the limited resources available. The opportunity costs incurred in practice depend upon the marginal changes of what is funded when a new intervention is adopted. Hence a suitable basis for CET measurement is the health gains produced by an additional dollar of health care spending, also called the 'marginal productivity' of health care spending - the CET can then be expressed as the reciprocal (costs/unit of health gain) of the marginal productivity.

This approach to determining the CET to guide decisions requires explicit consideration of the 'supply side' ability of the health care system to generate health gains. It is grounded in the realities of the conditions and constraints

prevailing in specific health care systems. A focus on the supply side is required in order to meet the ultimate objective of maximizing population health benefits from within the resources available.

Frequently, in applied CEAs intended to inform resource allocations in low- and middle-income countries, the choice of CETs has not been conceived in this way. Instead, CETs have often represented expressions of value – by some party (individuals, international organizations; though often undefined) – without consideration of the constraints of health care systems.

One such example is the World Health Organization recommended CETs of 1 to 3 times gross domestic product per capita (GDP pc) in a country. Another is CETs based upon individuals' statements about willingness to pay to improve their own health. This could be termed the 'consumption value of health'. It differs notably from individuals' willingness to contribute to collective health care funding (to improve the health of others) or Governments' abilities to generate health care resources. Such CETs are inherently 'demand-side' concepts.

The danger of using CETs conceived only from the demand-side is there is no guarantee they will reflect opportunity costs. If they are set too low, CEAs may recommend interventions are not funded when they could generate population health benefits. In this case there would be lost opportunities for health improvement. Alternatively, and more likely, if they are set too high, interventions could be recommended and funded that displace greater health benefits than they generate. Where evaluations focus on new and higher cost interventions, it is also likely these recommendations will exacerbate existing health inequalities.

The challenge for policy-makers, budget-holders and analysts alike is to determine and use CETs that reflect supply-side constraints. Unfortunately, however, there are few empirically estimated supply-side CETs. One exception is a study by Claxton et al (2015) that estimated the marginal productivity of the United Kingdom (UK) National Health Service (NHS) and produced a best estimate of the supply-side CET of £12,936, about half of UK GDP pc.

Two papers – Woods et al (2015) and Ochalek et al (forthcoming) - have attempted to estimate supply-side CETs for a wide range of countries. Woods et al (2015) extrapolate from the UK CET to estimate CETs for other countries in the world, based upon data on the relationship between country income and the willingness to pay for health gains. Ochalek et al (forthcoming) estimate CETs using published analyses estimating the causal effects of changes in health care spending on mortality and country-specific data to link this to other health outcomes. In both studies – although uncertainty bounds are wide – the estimated CETs are far lower than those previously posited by WHO: almost all are below 0.6 GDP p.c in a country; see Section 4 for full results.

The funding channels for healthcare delivery in many countries are complex and, where budget silos exist, CETs for particular programmes (e.g. where donor funding is for that programme alone) may differ from CETs estimated based

upon the marginal productivity of overall healthcare spending. However, this is an indication that allocations between disease programmes are unlikely to be optimal and, although at times there may be good reason for vertical funding, resources could be reallocated to where marginal productivity is highest to improve population health.

It should be highlighted that the estimates of Woods et al. and Ochalek et al. represent a starting point in the effort to derive empirically founded inputs to understanding and estimating true opportunity costs. Notable uncertainties in estimating the correct empirical supply-side CETs in low and middle-income countries persist (see Nakamura et al, forthcoming) and further research in this area would be very valuable. The estimates can be used to inform health care spending decisions but estimates from within countries (rather than relying on cross-country estimation) would help to corroborate or provide basis for revision of these estimates. In all cases, policymakers, budget holders and analysts should carefully consider the other claims on limited resources as well as their likely opportunity costs when informing or making resource allocation decisions.

Part 1: Introduction

Countries around the world have to make difficult decisions as to how healthcare is financed and scarce available resources are allocated to meet the health needs of their populations. The tools of economic evaluation can help to inform resource allocation decisions based upon comparison of the costs and consequences of alternative policy choices. Invariably, whether in Switzerland or Swaziland, the range of interventions that could offer health benefits to patients is beyond what can feasibly be funded from within the available resources. Choices therefore need to be made on the basis of the costs and benefits associated with alternative healthcare interventions and the *opportunity costs* of committing resources to those interventions – in terms of what benefits those resources could generate if used for alternatives priorities.

Economic evaluation is a set of tools that can be used to inform which interventions should be funded. A common approach to economic evaluation is to use incremental cost-effectiveness analysis (CEA) that compares the incremental costs (Δcosts) and incremental health benefits (Δhealth) of an intervention to other comparators. Results can then be expressed as an incremental cost effectiveness ratio (ICER) (Δcost/Δhealth); giving the cost-per-unit of health gain provided by the intervention. Health benefits are often represented in the form of quality adjusted life years (QALYs) gained or disability adjusted life years (DALYs) averted. To determine whether an ICER offers value for money, and can be justifiably deemed 'cost-effective', requires comparison to the opportunity costs of what would have to be given up if the intervention were funded. These are typically represented by a cost-effectiveness threshold (CET).

Notable research efforts have been committed to better understanding and measuring the health benefits of alternative interventions in low and middle income country settings, and similarly although less so their associated costs. In contrast there have been few studies to inform the choice of CETs – even though this is a critical component of cost-effectiveness assessments. The danger of applying a CET that is too low for any particular context is that an intervention may not be adopted that offers population health gains; whereas, and perhaps more likely, applying a CET that is too high risks interventions being adopted that displace/forgo more health gains than they generate. The choice of CET is therefore critical for efficient allocation of scarce healthcare resources. The choice ultimately determines which interventions are provided, to whom in the population, and who goes without needed – indeed often lifesaving – healthcare.

This report informs policy-makers (Part 2) and analysts/economic evaluation practitioners (Part 3) on alternative conceptualizations for CETs, and the assumptions underpinning, and implications, of these conceptualizations; and provides estimates of CETS that can be used in applied studies in a wide range of jurisdictions, for different kinds of decisions. Recommendations for policymakers and analysts/practitioners on the interpretation and presentation of findings from CEA studies are then presented (Part 4).

Part 2: Deciding when to invest in health care interventions - A guide for policymakers

Who is this guide for?

This guide (Part 2 of the report) is intended to inform budget-holders and policy-makers in a variety of institutions whose decisions have implications for healthcare resource allocation in low- and middle-income countries. Such institutions, and the types of decisions with which they are faced, include the following:

- Ministry of Health budget and planning departments – deciding how much of their available budgets to allocate to various departments, disease programmes, providers (e.g. central hospitals, regions, districts)

- Ministry of Health and other partner institutions – deciding upon what interventions to include in basic or essential healthcare packages

- Disease programmes (e.g. a HIV programme) – allocating their budgets across various technologies, drugs and interventions within their domain

- Social health and private insurance schemes – determining the interventions for which enrollees will be covered or reimbursed

- Health technology assessment (HTA) agencies – assessing which drugs, technologies and interventions should be funded by the public payer

- International funders (donors) or non-governmental organizations – deciding what types of healthcare provision to fund; either through domestic institutions or vertically (e.g. through their own or other non-Governmental providers)

- International funders and NGOs – deciding how to allocate their resources across countries.

The decisions made therefore relate to investments in different types of drugs, technologies, programmes or activities (herein referred to collectively as 'interventions') by a range of institutions for which a central objective can be regarded as improvement in population health.

What is the challenge facing policy-makers and budget holders?

Distinction can be made between revenue generation and pooling for healthcare provision, and the allocation of prepaid and pooled available resources across competing priorities. Whereas healthcare systems differ in how these primary

roles are carried out; particularly the extent to which collective healthcare is funded and the role of governments as single payers in the system; ultimately all healthcare systems exist around a central goal to bring about improvements in population health. The central resource allocation challenge then is to respond to the question 'How should available resources be allocated so as to maximize health improvement in the population?'. Although individual and collective values may lead some healthcare systems to also emphasize additional objectives (e.g. financial protection and equity-related concerns; see below), it is a reasonable assertion that health improvement is central to all.

Policymakers and budget holders may sit in different institutions and face different kinds of decisions (as shown above). Insofar as their primary constituency includes that population of a jurisdiction with entitlement to healthcare provision - determined through some mandated political process – all decision-makers should have responsibility towards and should be accountable to the population affected by their decisions; including those who benefit from their decisions and those who lose out. Given the primacy of public-sector payers in most healthcare systems, a useful benchmark to assess the value for money of any healthcare expenditure is therefore the health gains that could be attained with a change in the collective public healthcare budget (i.e. the marginal productivity of public healthcare spending).

A useful comparison for institutions that are not the primary organization involved in the delivery of healthcare in a country, is to assess whether the interventions/programmes they are considering funding would have greater or lower marginal productivity (i.e. generate more or fewer health gains per dollar spend) than if the funds were channeled into general public sector healthcare provision. If marginal productivity is higher, there may be a good reason to have or maintain vertical funding. However, if this is not the case, vertical funding is unlikely to be the optimal way to generate health gains and questions may be raised about the sustainability of such funding arrangements. Similarly, when the funding decision facing an organization (e.g. an international donor) is which jurisdiction to commit additional resources to, then the marginal productivities of different jurisdictions can be compared to assess in which locality additional spending would likely generate greatest health improvement.

What assessments are required when choosing a cost-effectiveness threshold to use within a jurisdiction or by an organization?

Good resource allocation decisions are those that best meet agreed social objectives e.g. using currently available resources. In healthcare, where the central objective is improvement in 'population health', a good decision should involve comparing the additional health benefits of an intervention with the health likely to be lost elsewhere as a consequence of any additional costs. This should also therefore be the aim of cost-effectiveness analysis and other forms of economic evaluation. To be consistent with actions likely to lead to population health improvement, a cost-effectiveness threshold should reflect these health

opportunity costs resulting from 'supply side' constraints in the ability of a healthcare system to generate health improvement.[1]

A variety of methods may be used to determine an appropriate CET that reflects heath opportunity costs resulting from supply-side constraints. In essence, the aim is to have reasonable understanding of the 'marginal productivity' of the healthcare system. This tells us how much health is expected to be generated (additional QALYs gained/DALYs averted) for a given additional expenditure, at the margin, and hence, how much health is expected to be foregone if the resource in question is spent on a new intervention. The possible methods and accompanying available estimates of CETs for use in different jurisdictions are presented in the next section. However, it is first useful to contrast these methods with some other ('demand-side') bases for CETs which have been used and which do not necessarily reflect opportunity costs.

Contrasting demand-side and supply-side estimates of CETs

Health opportunity costs, which represent the most suitable basis to determine CETs for use in resource allocation decisions related to constrained healthcare budgets, can be contrasted with opportunity costs that fall on other forms of consumption. These can include private consumption and consumption related to other forms of collective social expenditures (e.g. funded through tax receipts) for which the primary purpose is societal objectives other than health; such as government expenditures in other sectors including education or policing.

Particular confusion has arisen in the appropriate basis for estimating CETs as CETs applied in practice have typically been based upon individuals expressions of willingness to pay to improve (or prevent reduction of) their own health – so called 'willingness to pay' for health improvement.

It should be noted that individuals' willingness to pay to improve their own health may be different from their willingness to pay into collective healthcare spending (for which others in the population also benefit).It may also not reflect the available budgets – and the spending opportunities associated with these budgets – that result from the government/public payers' ability to generate and pool collective healthcare resources. Willingness to pay is a 'demand-side' approach to valuing health whereas a 'supply-side' concept is required to reflect what the real healthcare system is producing from the real budget constraint.

However, CETs have not generally been set to reflect health opportunity costs resulting from limited available healthcare budgets. For instance, values of GB£20-30,000 and US$50,000 per QALY have commonly been applied in the United Kingdom and United States, respectively; without clear rational but with some sense they reflect the consumption value of health. In low and middle income countries, the World Health Organization (WHO) has recommended thresholds of 1 to 3 times gross domestic product (GDP) per capita – seemingly on the basis of recommendations from the "Commission on Macroeconomics and Health" report from 2001.

It is not very clear what the basis of the 1-3 times GDP CETs really then is and, rather than representing any real notion of value they may simply be expressions of aspiration. However, the consequence of their use if they are inappropriate measures of the health opportunity cost is they are likely to reduce, rather than increases, population health. Supply-side measures of CETs are instead required.

What estimates of suitable thresholds for particular jurisdictions or organizations are available?

There are still worryingly few estimates of supply-side cost-effectiveness thresholds that exist at all across different jurisdictions. Although interest in estimating specific supply-side CETs has increased only very recently, to date just one study in one country - the UK – is available. This study provides an estimated CET based upon thorough analysis of country programmatic spending data and resulting health outcomes. Claxton et al. (2015)[1] estimated a causal link between changes in expenditure and mortality outcomes using data on different disease areas (programme budget categories). Additional information about the age and gender of the patient population was used to get from mortality to survival effects, and finally, data on health related quality of life (HRQoL) norms by age, gender and disease were used to obtain morbidity effects.

The research resulted in an estimated CET of £12,936 per QALY for 2008/9. It can be noted this is well below 1 times GDP per capita, which was around £26,000 at that time,.

A recent study, Woods et al. (2015)[2], uses the Claxton et al (2015) estimate and extrapolates to other countries using international income elasticities of the value of health. The approach relies upon some core assumptions set out fully in the paper, but it is not clear a priori whether these would likely lead to over- or under-estimates of CETs.

Woods et al. show opportunity costs are likely to be particularly high, and CETs low as a proportion of GDP per capita, in lower income countries. For instance the range of CET estimates for Ethiopia, a country with a 2013 GDP pc of US$505, is $10-$255 (2%-50% GDP pc); and for Indonesia, with a 2013 GDP pc of US$3,457, is $472-$1786 (14%-51% GDP pc). Again, this is in stark contrast to the WHO guidance.

Ochalek et al (forthcoming) employ a different approach, and provide a framework for generating country-level CETs using existing published cross-country estimates of the mortality effect of health expenditure.[3] Two different estimation strategies are used based upon published literature (Bohkari et al (2007)[4] and Moreno-Serra and Smith (2015)[5]). These are expanded upon using measures of mortality, survival and disability outcomes, reflecting the demographic and other characteristics of each LMIC.

The results also suggest that CETs representing likely health opportunity costs tend to be well below national GDP per capita in a large range of countries, although specific estimates of CETs depend upon the particular countries. For instance, using year 2000 values, CET estimates in Ethiopia are $6-$96 (3%-52% of GDP pc=$179 in 2000); and in Indonesia are $42-$244 (8%-47% of 2000 GDP pc=US$519 in 2000).

Supply-side CET estimates from the Woods et al. and Ochalek et al. studies for a full range of countries are provided in Section 4.

What CETs should be used if interventions draw upon resources not generally available for use across the whole health sector?

In many cases resources committed to particular interventions or to disease programmes (e.g. a country's HIV programme), come from sources other than domestic collectively funded healthcare budgets (e.g. international donors and non-governmental organizations). This funding is often "vertical" and therefore has to be spent on a specific intervention or programme. This funding does not therefore directly impose health opportunity costs on the country's limited budgets. In these cases, a suitable CET to inform decisions within such a budget silo may differ from the appropriate CET to inform the allocation of overall collective healthcare funds.

However, it should be noted that external funds could also be used for other purposes – in the same jurisdiction or even in other jurisdictions. In this sense they do incur health opportunity costs. For international funding decisions, donors can maximize the health gains resulting from their expenditures by committing resources to where the opportunity cost is likely to be highest (i.e. CETs lowest). The way in which external resource commitments are made requires scrutiny in this respect.

Are there other judgements, in addition to supply-side based CETs, that are required when deciding whether to invest in particular interventions?

Using supply-side CETs to inform resource allocation decisions assumes that all the benefits of interventions are health benefits and that all the costs fall on health care spending. However, other costs and consequences of interventions may also be relevant in health care decision making. They include wider impacts on families, communities, and other sectors of the economy (e.g. on educational outcomes). They may also include other (direct and indirect) costs (or savings) that are incurred in gaining access to an intervention or that result from associated health outcomes. For instance, these may include direct costs falling on individuals and families in accessing health interventions (e.g. travel, out-of-pocket and care costs), indirect time costs (e.g. loss of wages in individuals and informal carers), as well as costs falling on other sectors of the economy. Finally,

how health effects are distributed within a given jurisdiction, e.g. between the rich and the poor, could be of particular concern in many countries.

Non-health effects and costs that fall outside the health budget may be important because alternative interventions may result in different non-health effects that have social value. It is therefore useful, and even in some cases imperative, for policymakers to assess these when making decisions. In principle, non-health and societal impacts could be incorporated in a health economic evaluation. However, there remain methodological challenges in doing so, and it requires two things: firstly, knowledge of trade-offs between non-health and health benefits; and, secondly, knowledge and justification for determining who should make these trade-offs.

Deciding which non-health effects and which costs that fall outside the health budget should be included in primary analyses and who should trade-off health and non-health costs and benefits is therefore troublesome. Since there is no consensus on how to codify societal preferences, conflicts between different elements of social value may result. A particular concern is that health resources, primarily intended to generate 'health', may be used to meet other objectives that society may or may not deem to be as valuable as health itself.

Nevertheless, policymakers may wish to consider such non-health effects in the process of making their prioritization decisions.

Part 3: Informing health care investment decisions
- A guide for analysts

Who is this guide for?

This guide (Part 3 of the report) is intended for researchers and applied analysts working on economic evaluation studies to inform health care resource allocation decisions with implications in low and middle income countries. The possible users of such studies are those listed in Part 2.

A large number of economic evaluation studies are produced to guide decisions in low- and middle-income countries.[6] Some of these are well resourced and produced by analysts familiar with the leading and most technical of available methods. However, others are less well-resourced and need to guide decisions under much tighter time and financial constraints. The capacity for some analysts to keep pace with changes in methods may also be limited.

It is hoped that this guide will represent a first port of call for analysts looking for information on which cost-effectiveness thresholds they can use to inform decisions in their jurisdiction. It contains useful information for both well-resourced analyses and also studies operating within tighter resource and capacity constraints but nonetheless seeking to usefully inform policymaking. .

What types of decisions does this guide inform?

Cost-effectiveness thresholds (CETs) can inform investments in alternative clinical and other health interventions. They can also inform much larger shifts in healthcare funding, such as in the design of health benefits packages[7] or in changing the balance of funding between programmes (e.g. between hospital based and primary healthcare; or by disease programme).

Increasingly, economic evaluation studies are used to inform a wider range of decisions and, as central benchmarks of value in healthcare systems. CETs also have an important role for informing investments for which impacts on population health are less immediately clear. Examples include, the use of economic evaluation to inform investments in research and efforts to reduce decision uncertainty (e.g. through value of information analysis) or to strengthen healthcare systems and improve the uptake of clinical interventions (e.g. using value of implementation analysis).[8]

The important distinction between 'demand-side' and 'supply-side' CETs.

An important distinction must be made between alternative conceptual bases for CETs. In the past, many CETs used in applied analyses for low- and middle-income countries were detached from appropriate normative and theoretical bases for resource allocation in the context of constraints, and their impacts on population health outcomes were consequentially unclear.

Particular distinction can be made between 'demand-side' and 'supply-side' notions of CETs.

Various notions have been put forth to motivate the choice of CETs. In general these have been based upon expressions of the value of health (from various constituents – individuals, international organizations, doctors/experts). These value-based estimates are detached from an assessment of the capacity of healthcare systems to deliver interventions to a level that would be consistent with these expressions of value (i.e. they are 'demand-side' driven, without consideration of supply). In particular they are detached from constraints on the ability of health care systems to raise money.

In contrast, and more recently, greater attention has been paid to the necessity of grounding CETs in 'supply-side' assessments of the health benefits of competing calls on constrained healthcare budgets. In particular, where the objective is improvement in overall population health, an intervention should only be recommended for funding where a commitment of resources to that intervention will produce health benefits exceeding those displaced or forgone (i.e. opportunity costs) as a result of those resources becoming unavailable for the funding of other priorities.[9–12] To do otherwise, and apply demand-side CETs, risks both reducing population health and increasing health inequalities

(because interventions only likely to be accessible to a subsection of the population in need are likely to be prioritized).

It is therefore crucial for analysts to be aware of the conceptual/theoretical bases of alternative possible CETs they may encounter. Some CET estimates are grounded in demand-side notions of the value of health, and are inappropriate to inform allocation of constrained resources; whereas others (although currently, fewer) are based upon supply-side estimates of opportunity costs.

The following sections highlight widely encountered demand- and supply-side estimates and summarize the current state-of-knowledge on available supply-side estimates for use in a wide range of jurisdictions.

What 'demand side' CETs exist and have been used?

There are at least four widely encountered bases for CETs that predominantly rely on demand-side notions of value:

- Historical precedents of $100k and £20-30k per QALY thresholds applied in the United States and United Kingdom, respectively; and similar corresponding levels used in other countries.
- 1-3 times gross domestic product (GDP) per capita (pc) thresholds; previously recommended by the World Health Organization (WHO) and used in generalized cost-effectiveness analyses (GCEA) and other studies
- Stated preference elicitation studies of individuals' willingness to spend money to improve their own health (or reduce losses in health)
- Revealed preference elicitation studies of individuals' willingness to spend money to improve their own health (or reduce losses in health)

$100k and £30k per QALY CETs

In the United States, a CET of $50,000 per life year gained popularity in the 1990s. The true source of this benchmark is unknown and it was never endorsed by the Panel on Cost-Effectiveness in Health and Medicine who met in 1996.[13] However, in a meta-analytic review of end stage renal disease studies from 1968-1998, it was revealed this CET had morphed into one of $50k per QALY gained and was being widely applied in studies.[14] Gross (2008) concluded that the "$50,000 criterion is arbitrary and owes more to being a round number than to a well-formulated justification for a specific dollar value."[13]

Similarly, in the United Kingdom, a CET range of £20,000 to £30,000 per QALY has been used by the National Institute for Health and Care Excellence (NICE) since 2004. This range was used in decisions made prior to 2004, but is widely recognized (including by NICE) as a benchmark lacking empirical foundation.[15]

In summary, these values were based on precedent rather than having a clear scientific basis.

CETs 1-3 times GDP per capita in a country

Arguably the most well-known and widely applied 'demand side' CETs in low- and middle-income country are GDP p.c. based thresholds adopted by the World Health Organization for use alongside WHO-CHOICE.

Initially published in the "Commission on Macroeconomics and Health" report from 2001, the true origins of these CETs are unclear, but may have been derived from figures extrapolated using US based Value of a Statistical Live (VSL) of $6.3m from 1997. (See below: *Revealed preferences: the value of a statistical life studies*.) However, these estimates were intended to inform decisions regarding overall investments in healthcare. Their use of these thresholds when assessing the value of individual interventions from with constrained budgets is not consistent with population health improvement as they do not reflect the opportunity costs that are imposed on healthcare systems.

Stated preferences: social value of a QALY studies
The Health Intervention and Technology Assessment Program (HITAP) in Thailand, which was established in 2007, uses a CET based on estimates of willingness to pay (WTP) for health. This CET is intended to represent the "social valuation" of health versus other consumption and was estimated through stated preferences. This research, published in 2009, estimated the social value of a QALY through the assessment of utilities (through time trade off, TTO) and WTP. On the back of this work, the Health Economic Working Group under the Subcommittee for Development of the National List of Essential Drugs and the Subcommittee for Development of the Health Benefit Package and Service Delivery of the NHSO recommended a ceiling CET of 1x GDP per capita or 120,000 THB per QALY gained.[17] The organization has continued to evolve and conduct research, and the threshold was raised to 160,000 THB per QALY in 2013.[18] Although empirically derived, these estimates are aspirational, representing a notion of what ought to be, and should not be mistaken for 'supply side' CETs.

Revealed preferences: the value of a statistical life studies
CETs also exist which are based on value of a statistical life studies (VSL). VSL estimates can be derived through both revealed and stated preference studies. The former involves observing decisions relating to mortality risks and peoples willingness to pay to avoid risk, while the latter asks respondents to choose between hypothetical risk scenarios. Hirth et al (2000) meta-analysed various VSL estimates and used quality of life weights from the Beaver Dam Health Outcomes study to generate a QALY valuation for the US of $265,000 in 1997.[19] This is thought to be the basis of the estimates from the 2001 WHO Commission on Macroeconomics and Health report that forms the foundation of the 1-3 times GDP per capita CETs.

What 'supply side' thresholds exist and can be used?

There has been a paucity of supply-side estimates of cost-effectiveness thresholds (i.e. reflecting the marginal productivity of healthcare systems) in both high- and low-/middle-income countries settings alike. Currently, there

are only 3 known recent sources of supply-side CET estimates that can be used directly in economics evaluations:

- Claxton et al (2015); an estimate of the marginal productivity of the UK National Health Service (NHS)
- Woods et al. (2015); which extrapolates this estimate to other countries
- Ochalek et al. (forthcoming); providing alternative supply-side estimates for a wide range of countries.

Claxton et al (2015)
Claxton et al (2015) provides as yet the only example of a supply side based CET, which it does for the UK.[1] The authors make use of the rich data available in the UK on expenditure and mortality outcomes in different disease areas (programme budget categories), as well as health related quality of life (HRQoL) norms by age and gender and HRQoL associated with different diseases. The study estimates the effect of changes in spending on mortality outcomes by exploiting area-level variation in these variables and employing an instrumental variable (IV) approach to control for endogeneity (e.g. the possibility that mortality determines health care expenditures as well as being improved by it). From mortality outcomes, the authors determined deaths averted, and using additional information about the age and gender of the patient population, they determined the survival effects of changes in spending. With additional available information about HRQoL norms by age and gender and HRQoL associated with different diseases, the authors were able to determine the morbidity effects of changes in expenditure. Using the preferred set of assumptions, including using the effect of expenditure on mortality as a surrogate for the effect of expenditure on morbidity, the authors estimated a cost effectiveness threshold for 2008/9 of £12,936 per QALY.

Woods et al (2015)
Woods et al. (2015) uses the Claxton et al (2015) estimate and extrapolates this to other countries using information from the literature relating country income to willingness to pay for mortality risk reductions.[2] The validity of this work hinges upon the validity of the UK estimate of a supply-side CET and the validity of previous work looking at the relationship between country income and the value of a statistical life. It also rests upon two assumptions: (i) that the ratio of the supply-side CET to the demand-side CET is constant across countries and (ii) that the relationship between country income and the value of a statistical life can be translated directly to the relationship between country income and the value of a QALY. They show that CETs based on opportunity costs are likely much lower than those often used in decision-making in LMICs.

*Ochalek et al (2015**)***
Ochalek et al (forthcoming) take advantage of recent developments in econometric methods to control for endogeneity in the estimation of the mortality effects of changes in expenditure when using use cross-country data.[3] Using Bokhari et al (2007)[4] and Moreno-Serra and Smith (2015)[5] to represent

two different approaches within the literature, the authors show how cross-country econometric models can be used as an input for calculating country-specific CETs through analysis of other health outcomes, use of additional data and explicit modeling assumptions. Bokhari et al. (2007), using a cross-section of 127 countries from 2000, model the role of donor funding explicitly and allow for the endogeneity of key inputs into the health production function. Moreno-Serra and Smith (2015) use a panel of 148 countries between 1995 and 2008 and an innovative econometric modeling strategy that accounts for reverse causality with panel data fixed effects to control for country-level heterogeneity. Using the framework they have developed and applying it to results from Bokhari et al (2007) and Moreno-Serra and Smith (2015), Ochalek et al (forthcoming) estimate a range of CETs for each country. They find that the upper estimate of the range for nearly every country falls below 3x GDP per capita, and is below 1x GDP per capita for the vast majority of countries.

Thus applying these generic "rules of thumb" can do more harm than good: When interventions with ICERs that below 1 times GDP per capita but above the true 'supply side' CET are implemented, they will displace more health than they generate, resulting in a net health loss.

Part 4: A summary of the evidence on supply-side cost-effectiveness thresholds

Table 1 presents estimates of cost-effectiveness thresholds for a selected range of countries from Woods et al. (2015; CETs presented in 2013 US$)) and Ochalek et al (forthcoming; CETs 2000 US$). Results from all countries from both studies are presented in the Appendix.

Country	GDP per capita, 2013	Woods et al (2015) threshold range, 2013 US$	Threshold as a % of GDP	GDP per capita, 2000	Ochalek et al (forthcoming) threshold range, 2000 US$	Threshold as a % of GDP
Brazil	$11208	$2393 - $7544	21% - 67%	$3064	$575 - $1809	19% - 59%
Ethiopia	$505	$10 - $255	2% - 50%	$179	$6 - $93	3% - 52%
India	$1499	$115 - $770	8% - 51%	$548	$27 - $214	5% - 39%
Indonesia	$3475	$472 - $1786	14% - 51%	$519	$42 - $244	8% - 47%
Kazakhstan	$13610	$4485 - $8018	33% - 59%	$733	$196 - $310	27% - 42%
Malawi	$226	$3 - $116	1% - 51%	$184	$20 - $207	11% - 112%
Nepal	$694	$22 - $357	3% - 51%	$252	$32 - $141	13% - 56%
Thailand	$5779	$1181 - $3943	20% - 68%	$1770	$486 - $805	27% - 45%
Vietnam	$1911	$144 - $982	8% - 51%	$417	$110 - $369	26% - 89%

Table 1: 'Supply side' cost-effectiveness threshold estimates for selected countries

It is clear that in all countries the most likely CETs based upon empirical evidence are well below the 1 to 3 times GDP pc CETs that have frequently been

applied to date in low- and middle-income countries. In fact, with only a few exceptions the upper-bound estimates are below 60% of GDP pc, particularly in those countries with the lowest levels of per capita income. The implication is that use of CETs to inform resource allocation that are above these levels will likely reduce overall population health and may well exacerbate health inequalities.

This pattern can also be seen graphically. In Figure 1, the estimates from Ochalek et al. are inflated by a countries GDP pc growth (GDP pc 2013/GDP pc 2000) for comparability with Woods et al. In all but a few cases the CET estimates are below the line of 1 times GDP pc capita; although uncertainty bounds are nevertheless wide.

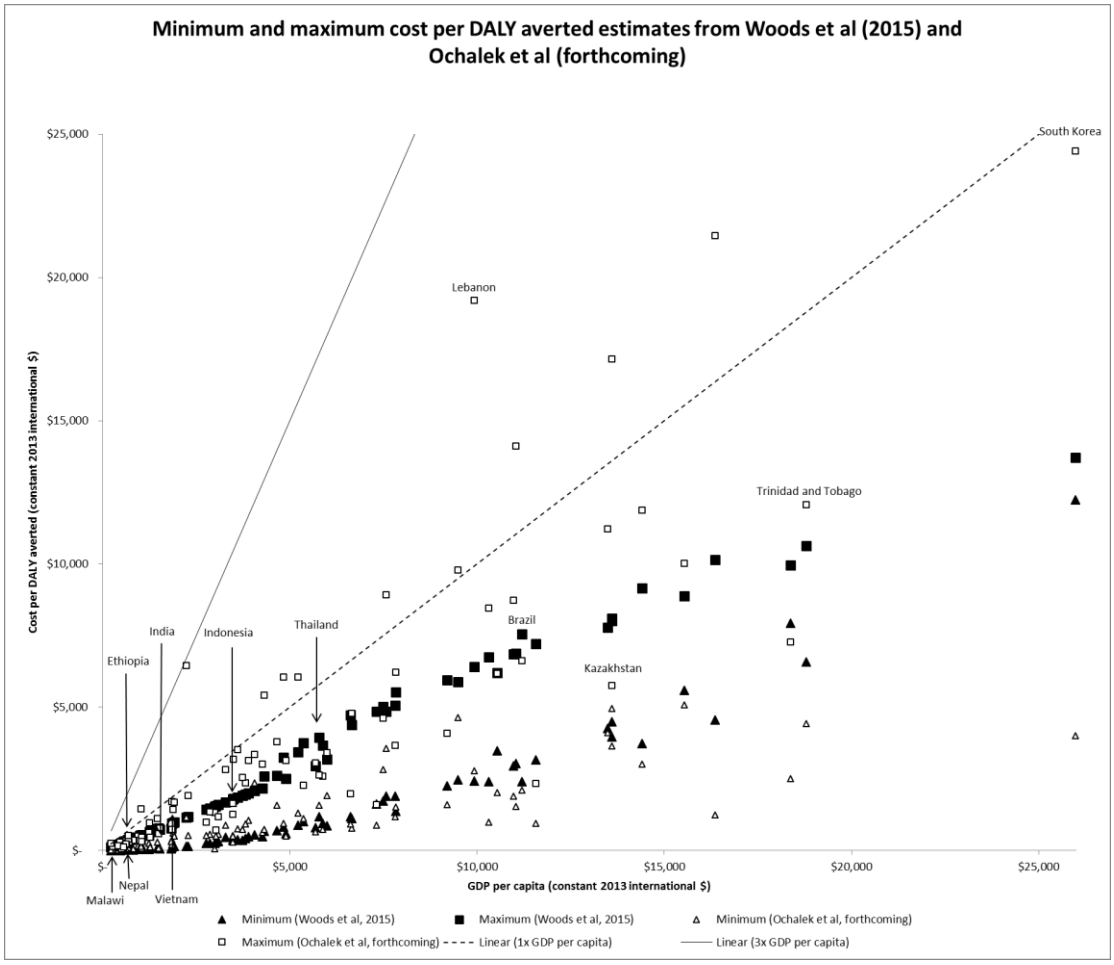

Figure 1

These estimates are in no way intended to provide 'final answers' to what supply-side CETs are appropriate for different countries. The estimates presented may be considered as plausible values and input to inform resource allocation decisions. More research in this area is urgently needed. This is further underlined by Nakamura et al., (forthcoming).

In their extensive analysis, Nakamura et al., re-examine the literature that identifies the impact of health care expenditures on mortality outcomes, using cross-country data. Employing exactly the same data and econometric

specifications as the two key published studies – Bokhari et al., (2007) and Moreno-Serra and Smith (2015), that also represent the backbone of the above approach by Ochalek et al., (2015) – Nakamura et al., start by successfully replicating the findings of the original studies. However, further analyses using updated data and more 'streamlined' econometric specifications, plus statistical data imputation and extensive robustness checks, reveal a considerable degree of sensitivity in the results.

These mixed findings should not be taken to imply that the conclusions from the previous studies (and – by implication – from the Ochalek et al., (2015) study) are necessarily invalid, as countries and survey years covered in the published data and in our updated data do indeed differ, and therefore it may not be a complete surprise that the analysis leads to different conclusions.

Nevertheless, our findings highlight the potential lack of generalizability of the results to different settings. They also underline the need for further improvement in either data, empirical methods, or even research design. A particularly promising future strategy could be to estimate opportunity costs using *within* country data, based on formal evaluations of existing policy changes within expanded government health care expenditure, i.e. a "natural experimental" approach.

Bibliography

1. Claxton, K. *et al.* Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technol. Assess.* **19,** 1–503, v–vi (2015).
2. Woods, B., Revill, P., Sculpher, M. & Claxton, K. *Country-level cost-effectiveness thresholds: initial estimates and the need for further research*. (2014).
3. Ochalek, J., Claxton, K. & Lomas, J. Cost per DALY thresholds for low and middle income countries: evidence from cross country data. *(forthcoming)*
4. Bokhari, F. A. S., Gai, Y. & Gottret, P. Government health expenditures and health outcomes. *Health Econ.* **16,** 257–73 (2007).
5. Moreno-Serra, R. & Smith, P. C. Broader health coverage is good for the nation's health: evidence from country level panel data. *J. R. Stat. Soc. Ser. A. Stat. Soc.* **178,** 101–124 (2015).
6. *Bill and Melinda Gates Foundation Methods for Economic Evaluation Project (MEEP) Final Report*. (2014). at <https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-International/projects/MEEP-report.pdf>
7. Sculpher, M. in *Design the Health Benefits Package* (Center for Global Development).
8. Fenwick, E., Claxton, K. & Sculpher, M. The value of implementation and the value of information: combined and uneven development. *Med. Decis. Making* **28,** 21–32
9. Revill P, Asaria M, Phillips A, Gibb DM, G. C. *WHO decides what is fair? International HIV treatment guidelines, social value judgements and equitable provision of lifesaving antiretroviral therapy*. (2014).
10. Newall, A. T., Jit, M. & Hutubessy, R. Are current cost-effectiveness thresholds for low- and middle-income countries useful? Examples from the world of vaccines. *Pharmacoeconomics* **32,** 525–31 (2014).
11. Marseille, E., Larson, B., Kazi, D., Kahn, J. & Rosen, S. WHO | Thresholds for the cost–effectiveness of interventions: alternative approaches. *Bull. World Health Organ.* **93,** 118–124 (2015).
12. Revill, P. *et al. Using cost-effectiveness thresholds to determine value for money in low- and middle-income country healthcare systems: are current international norms fit for purpose?* (2014). at <https://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP98_costeffectiveness_thresholds_value_low_middle_income_countries.pdf>
13. Grosse, S. D. Assessing cost-effectiveness in healthcare: history of the $50,000 per QALY threshold. *Expert Rev. Pharmacoecon. Outcomes Res.* **8,** 165–78 (2008).
14. Winkelmayer, W. C., Weinstein, M. C., Mittleman, M. A., Glynn, R. J. & Pliskin, J. S. Health economic evaluations: the special case of end-stage renal disease treatment. *Med. Decis. Making* **22,** 417–30
15. Claxton, K., Sculpher, M., Palmer, S. & Culyer, A. J. Causes for concern: is NICE failing to uphold its responsibilities to all NHS patients? *Health Econ.* **24,** 1–7 (2015).
16. Jirawattanapisal, T., Kingkaew, P., Lee, T.-J. & Yang, M.-C. Evidence-based

decision-making in Asia-Pacific with rapidly changing health-care systems: Thailand, South Korea, and Taiwan. *Value Health* **12 Suppl 3,** S4–11 (2009).
17. Shiroiwa, T. *et al.* International survey on willingness-to-pay (WTP) for one additional QALY gained: what is the threshold of cost effectiveness? *Health Econ.* **19,** 422–37 (2010).
18. Ryen, L. & Svensson, M. The willingness to pay for a quality adjusted life year: a review of the empirical literature. *Health Econ.* (2014). doi:10.1002/hec.3085
19. Hirth, R. A., Chernew, M. E., Miller, E., Fendrick, A. M. & Weissert, W. G. Willingness to pay for a quality-adjusted life year: in search of a standard. *Med. Decis. Making* **20,** 332–42

Appendix – Supply side cost-effectiveness threshold estimates for all countries

(*n.b. estimates for some countries are missing where national gross domestic product per capita estimates are missing)

Country	GDP per capita, 2013	Woods et al (2015) threshold range	Threshold as a % of GDP	GDP per capita, 2000	Ochalek et al (forthcoming) threshold range	Threshold as a % of GDP
Armenia	$3505	$387 - $1801	11% - 51%	$650	$305 - $591	47% - 91%
Azerbaijan	$7812	$1901 - $5051	24% - 65%	$476	$71 - $223	15% - 47%
Bangladesh	$958	$30 - $427	3% - 45%	$486	$78 - $255	16% - 52%
Benin	$805	$20 - $414	2% - 51%	$275	$25 - $108	9% - 39%
Burkina Faso	$684	$17 - $379	2% - 55%	$279	$25 - $82	9% - 29%
Burundi	$267	$3 - $137	1% - 51%	$137	$12 - $73	9% - 53%
Cambodia	$1007	$44 - $518	4% - 51%	$505	$58 - $176	12% - 35%
Chad	$1054	$31 - $540	3% - 51%	$183	$15 - $79	8% - 43%
Comoros	$815	$19 - $452	2% - 55%	$565	$47 - $241	8% - 43%
Congo, Dem. Rep.	$484	$5 - $230	1% - 47%	$677	$23 - $132	3% - 20%
Cote d'Ivoire	$1529	$61 - $737	4% - 48%	$437	$28 - $221	6% - 51%
Eritrea	$544	$9 - $280	2% - 51%	$127	$19 - $76	15% - 60%
Ethiopia	$505	$10 - $255	2% - 50%	$179	$6 - $93	3% - 52%
Gambia, The	$489	$12 - $252	2% - 52%	$863	$92 - $151	11% - 18%
Georgia	$3605	$366 - $1850	10% - 51%	$502	$106 - $491	21% - 98%
Ghana	$1858	$104 - $951	6% - 51%	$289	$32 - $114	11% - 39%
Guinea	$523	$9 - $269	2% - 51%	$747	$56 - $72	7% - 10%
Guinea-Bissau	$564	$9 - $256	2% - 45%	$218	$31 - $72	14% - 33%
India	$1499	$115 - $770	8% - 51%	$548	$27 - $214	5% - 39%
Indonesia	$3475	$472 - $1786	14% - 51%	$519	$42 - $244	8% - 47%
Kenya	$1246	$32 - $519	3% - 42%	$241	$48 - $120	20% - 50%
Kyrgyz Republic	$1263	$58 - $649	5% - 51%	$265	$105 - $202	39% - 76%
Madagascar	$463	$9 - $235	2% - 51%	$176	$24 - $111	14% - 63%
Malawi	$226	$3 - $116	1% - 51%	$184	$20 - $207	11% - 112%
Mali	$715	$17 - $368	2% - 51%	$174	$13 - $71	8% - 41%
Mauritania	$1069	$46 - $550	4% - 51%	$405	$54 - $121	13% - 30%
Moldova	$2239	$148 - $1151	7% - 51%	$247	$128 - $713	52% - 288%
Mongolia	$4056	$543 - $2085	13% - 51%	$209	$122 - $173	58% - 82%
Mozambique	$605	$8 - $294	1% - 49%	$449	$72 - $155	16% - 35%
Nepal	$694	$22 - $357	3% - 51%	$252	$32 - $141	13% - 56%
Nicaragua	$1851	$118 - $937	6% - 51%	$915	$185 - $842	20% - 92%
Niger	$415	$5 - $213	1% - 51%	$199	$8 - $77	4% - 39%
Nigeria	$3006	$239 - $1545	8% - 51%	$148	$3 - $65	2% - 44%
Pakistan	$1275	$87 - $669	7% - 52%	$368	$36 - $131	10% - 36%
Rwanda	$639	$13 - $323	2% - 51%	$379	$43 - $119	11% - 32%
Senegal	$1047	$34 - $544	3% - 52%	$465	$68 - $136	15% - 29%
Tajikistan	$1037	$37 - $533	4% - 51%	$118	$17 - $166	14% - 140%
Tanzania	$695	$18 - $357	3% - 51%	$136	$19 - $101	14% - 74%

Togo	$636	$13 - $327	2% - 51%	$368	$39 - $135	11% - 37%
Uganda	$572	$11 - $293	2% - 51%	$392	$49 - $88	12% - 22%
Ukraine	$3900	$487 - $2005	12% - 51%	$685	$185 - $552	27% - 81%
Uzbekistan	$1878	$138 - $965	7% - 51%	$436	$141 - $333	32% - 76%
Vietnam	$1911	$144 - $982	8% - 51%	$417	$110 - $369	26% - 89%
Yemen, Rep.	$1473	$83 - $757	6% - 51%	$143	$29 - $108	20% - 76%
Zambia	$1845	$68 - $768	4% - 42%	$159	$27 - $82	17% - 51%
Albania	$4659	$702 - $2612	15% - 56%	$1140	$386 - $931	34% - 82%
Algeria	$5361	$1012 - $3743	19% - 70%	$1197	$248 - $507	21% - 42%
Belarus	$7575	$1895 - $4857	25% - 64%	$1056	$497 - $1244	47% - 118%
Belize	$4894	$584 - $2503	12% - 51%	$3407	$354 - $2190	10% - 64%
Bolivia	$2868	$250 - $1474	9% - 51%	$686	$136 - $323	20% - 47%
Bulgaria	$7499	$1720 - $5025	23% - 67%	$1580	$594 - $973	38% - 62%
Colombia	$7831	$1370 - $5518	17% - 70%	$2298	$444 - $1827	19% - 80%
Dominican Republic	$5879	$937 - $3675	16% - 63%	$2702	$335 - $1194	12% - 44%
Ecuador	$6003	$858 - $3191	14% - 53%	$833	$265 - $474	32% - 57%
El Salvador	$3826	$422 - $1967	11% - 51%	$2086	$508 - $1286	24% - 62%
Guatemala	$3478	$360 - $1788	10% - 51%	$1363	$194 - $495	14% - 36%
Guyana	$3739	$348 - $1924	9% - 51%	$964	$191 - $657	20% - 68%
Honduras	$2291	$149 - $1177	7% - 51%	$1036	$235 - $868	23% - 84%
Jordan	$5214	$872 - $3432	17% - 66%	$1133	$284 - $1316	25% - 116%
Kazakhstan	$13610	$4485 - $8018	33% - 59%	$733	$196 - $310	27% - 42%
Lithuania	$15538	$5598 - $8886	36% - 57%	$3306	$1081 - $2134	33% - 65%
Macedonia, FYR	$4838	$824 - $3246	17% - 67%	$1973	$389 - $2468	20% - 125%
Morocco	$3093	$316 - $1590	10% - 51%	$1265	$231 - $478	18% - 38%
Namibia	$5693	$791 - $2958	14% - 52%	$2483	$279 - $1330	11% - 54%
Paraguay	$4265	$484 - $2179	11% - 51%	$1471	$216 - $1037	15% - 71%
Peru	$6662	$1114 - $4383	17% - 66%	$1812	$208 - $1300	11% - 72%
Philippines	$2765	$256 - $1421	9% - 51%	$1241	$235 - $447	19% - 36%
Romania	$9499	$2467 - $5875	26% - 62%	$1666	$813 - $1716	49% - 103%
Sri Lanka	$3280	$453 - $1686	14% - 51%	$757	$204 - $653	27% - 86%
Swaziland	$3034	$288 - $1559	9% - 51%	$1510	$250 - $352	17% - 23%
Thailand	$5779	$1181 - $3943	20% - 68%	$1770	$486 - $805	27% - 45%
Tunisia	$4317	$678 - $2592	16% - 60%	$2382	$400 - $2989	17% - 125%
Turkey	$10972	$2950 - $6861	27% - 63%	$2826	$487 - $2249	17% - 80%
Botswana	$7315	$1621 - $4839	22% - 66%	$3056	$365 - $667	12% - 22%
Brazil	$11208	$2393 - $7544	21% - 67%	$3064	$575 - $1809	19% - 59%
Croatia	$13608	$3953 - $8101	29% - 60%	$4040	$1474 - $5090	36% - 126%
Estonia	$18783	$6574 - $10636	35% - 57%	$4309	$1017 - $2770	24% - 64%
Gabon	$11571	$3164 - $7218	27% - 62%	$1820	$148 - $369	8% - 20%
Hungary	$13481	$4268 - $7773	32% - 58%	$4708	$1433 - $3921	30% - 83%
Korea, Rep.	$25977	$12227 - $13722	47% - 53%	$10051	$1548 - $9444	15% - 94%
Lebanon	$9928	$2420 - $6416	24% - 65%	$2297	$642 - $4441	28% - 193%
Malaysia	$10538	$3481 - $6192	33% - 59%	$2884	$552 - $1691	19% - 59%

Mauritius	$9203	$2248 - $5945	24% - 65%	$4146	$720 - $1847	17% - 45%
Mexico	$10307	$2410 - $6749	23% - 65%	$5748	$554 - $4716	10% - 82%
Panama	$11037	$3042 - $6869	28% - 62%	$3028	$422 - $3874	14% - 128%
South Africa	$6618	$1175 - $4714	18% - 71%	$3722	$519 - $1119	14% - 30%
Trinidad and Tobago	$18373	$7941 - $9959	43% - 54%	$4102	$561 - $1623	14% - 40%
Uruguay	$16351	$4548 - $10147	28% - 62%	$5966	$450 - $7827	8% - 131%
Venezuela, RB	$14415	$3724 - $9151	26% - 63%	$2362	$493 - $1945	21% - 82%

Supplementary Table 1: Results from all countries from Woods et al (2015) and Ochalek et al (forthcoming)