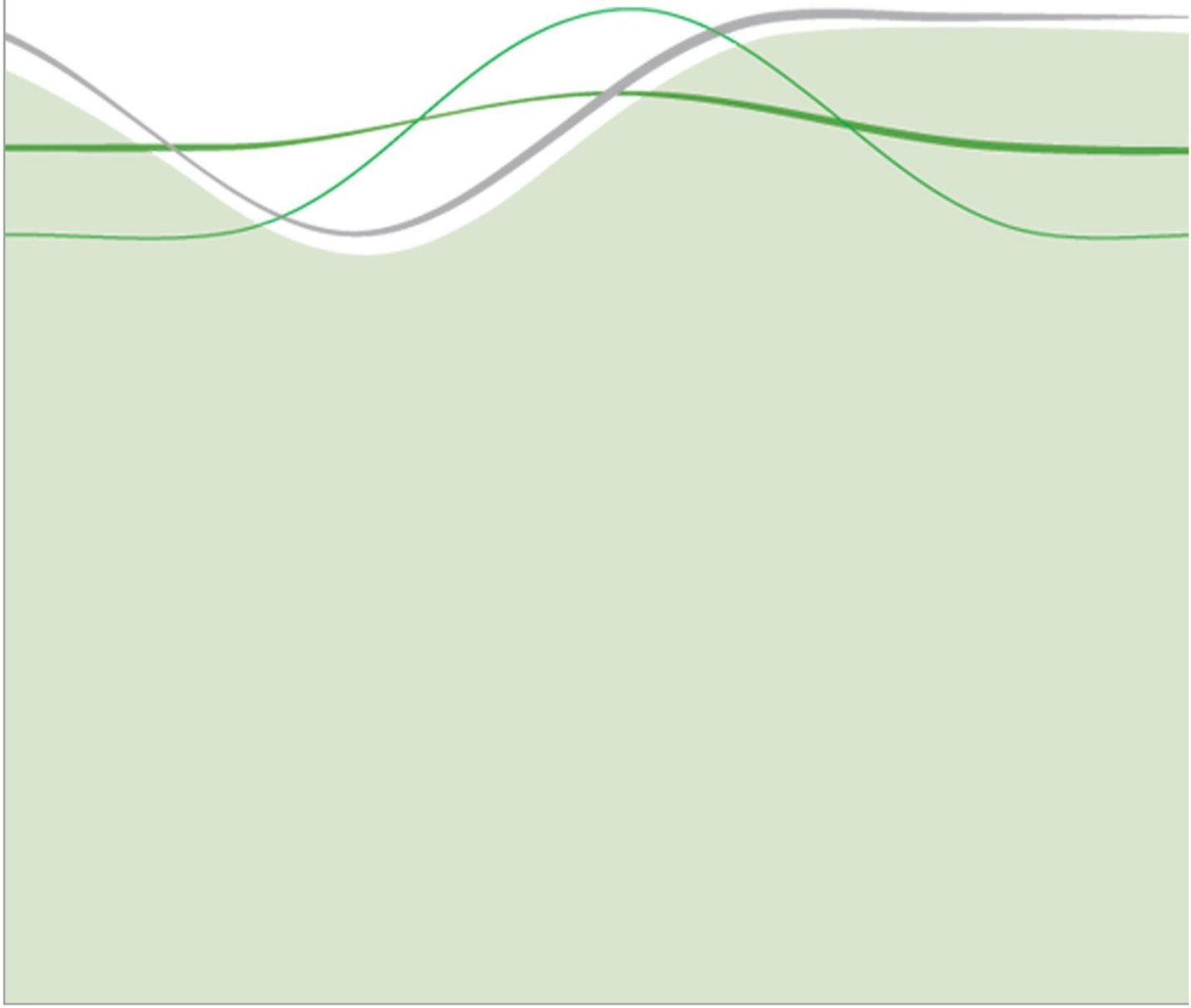
Office of Health Economics Research

Research Paper 18/07

Issues Surrounding the Estimation of the Opportunity Cost of Adopting a New Health Care Technology: Areas for Further Research

August 2018

Karla Hernandez-Villafuerte, Bernarda Zamora, and Adrian Towse



Issues Surrounding the Estimation of the Opportunity Cost of Adopting a New Health Care Technology: Areas For Further Research

Karla Hernandez-Villafuerte[a], Bernarda Zamora[b] and Adrian Towse[b]

[a]German Cancer Research Center (DKFZ)

[b]Office of Health Economics

August 2018

Please cite this report as:

Hernandez-Villafuerte, K., Zamora, B. & Towse, A. (2018). R&D, Issues Surrounding the Estimation of the Opportunity Cost of Adopting a New Health Care Technology: Areas For Further Research. OHE Research Paper 18/07, London: Office of Health Economics. Available at: https://www.ohe.org/publications/issues-surrounding-estimation-opportunity-cost-adopting-new-health-care-technology

The Office of Health Economics (a company limited by guarantee of registered number 09848965; a charity with registered charity number: 1170829)
Southside, 7th Floor
105 Victoria Street
London SW1E 6QT
United Kingdom

For further information please contact:
Adrian Towse
Director
atowse@ohe.org
Tel: +44 207 747 8858

About OHE Research Papers

OHE Research Papers are intended to provide information on and encourage discussion about a topic in advance of formal publication. Any views expressed are those of the authors and do not necessarily reflect the views or approval of OHE, its Editorial Panel or Research and Policy Committee, or its sponsors.

Once a version of the Research Paper's content is published in a peer review journal, that supersedes the Research Paper and readers are invited to cite the published version in preference to the original version.

Funding and Acknowledgements

OHE received research funding from the Association of the British Pharmaceutical Industry (ABPI) towards the preparation of this paper. OHE retained editorial control throughout and the content of this paper is the sole responsibility of the authors.

Jon Sussex of RAND Europe and the Cambridge Centre for Health Services Research (CCHSR) worked with KVH, when he was at the Office of Health Economics, on the analysis which is summarised in the Annex of the paper**.**

The methodological approach set out in the paper was developed by KVH. KVH prepared the first draft of the paper. This was edited and added to by AT and BZ. AT is an author of an earlier paper, Barnsley et al. (2013), on which this paper's research agenda builds.

Table of Contents

Abstract

Funding new medical technologies has opportunity costs. From the supply-side perspective of a tax-funded health care system such as the English NHS, opportunity costs reflect health losses due to displaced medical technologies given a fixed budget constraint. Empirical methods used to date to estimate the threshold assume observed health outcomes and health budgets reveal health system preferences. The threshold has been estimated from the average cross-sectional change in mortality as predicted by changes in the health budget, with assumptions made for quality of life gains.

In this paper we set out a research agenda proposing alternative approaches in three areas to improve understanding of supply side opportunity costs for the English NHS. A subsequent paper will set out our findings.

First, it is possible that some purchasers can introduce health services by improving efficiency rather than displacing an existing activity. We set out how Data Envelopment Analysis (DEA) can be used to explore differences in commissioner efficiency.

Second, we expect that health authorities (PCTs) will have differences in their production functions. We set out evidence of this from a cluster analysis of PCTs. If so, purchasers will displace those services with relatively smaller health losses. Current approaches estimate the relationship between expenditures and health outcomes using a linear regression methodology. However, if decisions made as a result of the approval of a new health technology are considered at the margin and not at the mean, and the health locations are heterogeneous, a methodology that does not focus on the mean of the health outcome is preferable. One possible approach is to employ a quantile regression function which can accommodate non-linearity in the relationship between expenditures and outcomes, and variation in outcome elasticities at the margin. This can be done by estimating models for several conditional quantile functions.

Thirdly, there are observable health outcomes reflecting purchaser priorities other than reducing mortality. Use of DEA to analyse multi-outcome decisions will allow a more robust estimate of the cost effectiveness threshold given the sources of local heterogeneity.

Our proposed methods aim at gaining insight as to extent of the heterogeneity of supply side opportunity costs across English health services and how an estimate of a supply side threshold that takes account of this might be established.

1. INTRODUCTION

The idea of using an opportunity cost threshold for the efficient allocation of resources across different treatment alternatives is not new. It has been discussed for 45 years (Weinstein and Zeckhauser 1973). As the value for money for interventions came to be expressed as incremental cost effectiveness ratios (ICERs), following Weinstein and Stason (Weinstein and Stason 1977), interest developed in using an ICER threshold as a decision rule that determined what is and is not considered acceptable value for money by a health system. There is, however, no consensus as to which methods to use to estimate an evidence-based ratio for health care policy makers. This is not because of lack of interest. A threshold is considered important for decisions in a number of health systems. Countries such as the UK, Canada and Australia apply formal and informal thresholds to decide on the adoption of new health treatments (George, Harris, and Mitchell 2001; Schwarzer et al. 2015). There appears to be no empirical evidence supporting the thresholds that are implicitly or explicitly used. This maybe because the estimation of a threshold is not a trivial exercise.

1.1. "Demand side" and "supply side" concepts of the threshold

There exist different conceptualisations of what the threshold should reflect. These can be divided into two, with nuances of meaning inside each classification: "demand" and "supply" side approaches to estimating opportunity cost. The former seeks the societal monetary valuation by its citizens of, or willingness to pay for, health gains. The opportunity cost is what the citizen as social or private insurance premium payer or as tax payer is willing to give up in spending on other goods and services in order to buy more health gain for themselves or their fellow citizens. The latter considers the threshold as an estimate of the health losses caused by the reallocation of resources needed to fund a new treatment given a fixed health care system budget constraint. These health losses are the opportunity cost of a new treatment (Cleemput et al. 2011).

Much of the recent focus on generating an evidence-based threshold has been focussed on this "supply" side view of opportunity cost. Different approaches have been tested to uncover a meaningful supply side threshold value (Vallejo-Torres et al. 2016). There are multiple reasons for a lack of consensus about both appropriate methods to calculate a supply side estimate, and the resulting value obtained from a particular method, given the number of assumptions that have to be used.

In a health system with a tax-funded National Health Service, such as the UK, it is to be expected that more attention is given to a threshold based on the "supply side" opportunity cost of adopting a new medical technology, given a fixed budget constraint. In the UK, the annual NHS budget is set in advance (and approved by Parliament) and can be considered fixed at a national level in any one year. This means that the money that needs to be efficiently allocated is, to some extent, fixed. Any additional health intervention might be expected, in theory, to be linked to a displacement of some other activity. This displacement represents the immediate opportunity cost and can also be referred to as the shadow price of the new intervention (Vallejo-Torres et al. 2016; Cleemput et al. 2011). An additional argument used in favour of the supply side estimated threshold in the UK is that societal preferences for spending appear to have no clear relationship to the process of defining the total NHS budget. The adoption of new technologies based on a demand side estimated threshold (societal willingness to pay) could lead to increases in expenditure that exceed the limited budget (Culyer et al. 2007). This is supported by the finding that demand side estimated thresholds are

considerably higher than supply side estimated thresholds (Vallejo-Torres et al. 2016). One implication of this is that health systems may be underfunded as compared to citizens' willingness to pay for health gain. Another is that the main focus of work in estimating the threshold has been on supply side estimates on the assumption that this is the binding constraint.

With this in mind, researchers have attempted to estimate an evidence-based threshold value reflecting the supply side opportunity cost to the health system using revealed preference techniques. Only a few approaches have been considered, reflecting the difficulties of estimating opportunity cost. These fit into one of two categories.

1.2. Two approaches to estimating the supply side threshold

1.2.1. Observing / interviewing purchasers about their decisions

One approach to estimating the threshold is by looking at the cost-effectiveness of those health care services introduced or discontinued over a period of time. If local health authorities (purchasers/commissioners) are making decisions considering their opportunity cost, they will adopt services and treatments below the threshold and exclude those above, which permits the approximation of the threshold. This conjecture was tested by Appleby et al. (2009). The researchers found, however, the few purchaser decisions were informed by an understanding of their cost-effectiveness, which hampered the ability to approximate an implicit threshold value.

1.2.2. Estimating the relationship between health outcomes and health care expenditures

A second approach to approximate the health losses of adopting a new treatment is based on the estimation of the relationship between health outcomes and health care expenditures. Lichtenberg (2004), in one of the first estimations of the supply side threshold, used time series analysis (1960-2001) to understand the relationship between life expectancy and health expenditures in the case of the USA healthcare system, considering previous values of life expectancy and pharmacological innovation (new molecular entities) as an approximation of technological change. Claxton et al. (2013), Claxton et al. (2015)[1] take the approach further by examining the relationship between expenditure and outcomes using a cross-sectional analysis of data from English health locations. They derived an England specific opportunity cost based on separate estimations for each of the 23 clinical areas (Programme Budget Category, PBC). The Claxton et al. (2013) analysis was based on the methods previously applied by Martin, Rice, and Smith (2008, 2012) where the threshold values for particular clinical areas, including cancer and circulatory problems, were estimated.

Vallejo-Torres, García-Lorenzo, and Serrano-Aguilar (2016), using cross-sectional and time series data on expenditures and health outcomes of 17 different regional health services in Spain, evaluated what they termed "a health outcome elasticity to changes in expenditures" which can be regarded as a threshold.

[1] Claxton et al. (2013) is the Centre for Health Economics Working Paper and Claxton et al. (2015) is the journal publication. For the authors ease of reference we use the (2013) version as the reference in this paper.

None of these estimates is currently used by policy makers, although the DH in England now uses a £15,000 per QALY threshold based on the Claxton et al. (2013) work[2] for some purposes.

2. OUR APPROACH

In our view there is a need for additional approaches that build on the progress made by researchers in this topic. It is important to understand what can be considered a relevant (evidence-based) threshold and why previous estimations of the threshold value may not fully capture this. The aim of this paper is to explore implicit and explicit assumptions that hinder the correct estimation of the threshold. We focus only on the supply side estimate of the threshold: the opportunity cost acceptable to adopt a new health treatment, which measures the health loss resulting from services displaced by purchasers of health care when a new medical technology is adopted.

In order to be able to make recommendations that improve on the currently available methodologies, three questions are explored. We consider them in turn.

2.1. Are purchasers always displacing services when a new medical technology is adopted?

There are two alternative possibilities. Firstly, displacement occurs under fixed budgets, yet budgets are not entirely fixed at a particular health location (e.g. a local health authority); purchasers that underspend might choose to underspend less; those that overspend might choose to overspend more (Sussex and Hernandez-Villafuerte 2015).

Secondly, recent empirical studies of the NHS in different parts of the UK suggest that one of the first responses to the approval of a new medical technology by local health care purchasers is to squeeze greater efficiency out of health care providers, such that the opportunity cost is not reflected in the pound-for-pound displacement of other health services they reimburse (Karlsberg Schaffer et al. 2015; Karlsberg Schaffer et al. 2016). In spite of these additional possibilities, approaches to estimate the threshold implicitly assume that there is always a displacement involving a loss of health. If displacement is not the only option, an opportunity cost estimated on the basis of that it is will be greater than the actual opportunity cost.

2.2. If there is displacement, what services are displaced and where?

To be sure about the appropriate threshold we need to know which health services purchasers are giving up to introduce a new medical technology. The results of the analysis conducted by Karlsberg Schaffer et al. (2016) suggested that linking NICE Technology Assessments (TAs) to specific service displacements is very difficult. Decisions are not informed by explicit prioritisation activities. Current estimates of the threshold bypass this lack of information by averaging the effects of changes in expenditure by clinical area (e.g. Programme Budget Categories) and estimating the QALY gain that resulted from that expenditure. There are two main reasons why there may be problems with such an approach.

[2] £15,000 per QALY is the recommended threshold according to Department of Health internal guidance. However, there is no public domain policy statement setting out the £15,000 number.

Firstly, observed differences in the threshold across clinical areas may reflect one or more of the following factors:

- different priorities across disease areas within the health system – so even though it may be more difficult (easier) to generate health gain in an area, its importance (lack of importance) means that more (less) is spent on it;
- different health outcomes depending on the clinical area. If non-QALY objectives are important, then estimates of a cost-per-QALY threshold will vary because the QALY is not fully reflecting the outcomes purchasers are seeking. A different mix of outcomes between disease areas, even as between mortality and morbidity objectives, will also complicate accurate estimation of a composite threshold measure that is valid across clinical areas;
- inefficiencies in some clinical areas. For example, some priority areas may be less efficient as there is less pressure to control costs. This is over and above any willingness to pay for technologies that deliver health gain at a higher cost.
- the difficulty of measuring the marginal costs of different Programme Budget Categories as costs are allocated between Categories based on accounting rules of thumb.

Secondly, purchasers in diverse health locations could react differently by clinical area to a requirement to adopt a new medical technology. Estimating a threshold reflecting the average displacement across all health locations does not take account of the likelihood that in a national health system where there are variations as between health locations in budget, variations in costs arising from different local market conditions for inputs, different technical constraints, as well as differences in the relative importance they attach to particular clinical areas. It also might be expected that health locations will displace services on those clinical areas with lower health losses. If so, the displacement in a given health location could be smaller than suggested by the national average of the relationship between expenditures and health outcome per clinical area. Which clinical areas have the lowest opportunity cost will depend on the characteristics of the health locations and their providers. The estimation of a single threshold representing an "average health location" may not be a good guide to displacement across clinical areas and health locations.

2.3. What health outcomes reflect the health system priorities?

The definition of health outcome employed to reflect the change in expenditures is one of the most important assumption in the determination of the threshold value. Claxton et al. (2013) used QALYs, derived indirectly, as the key outcome measure and estimated these by multiplying improvements in mortality by the ratio of mortality and morbidity DALYs in the global burden of disease calculations for England, given that no direct estimates for morbidity (quality of life) improvements are available. Lichtenberg (2004) based his analysis on life expectancy at birth. Health outcomes in a number of clinical areas may not be well reflected solely in improvements in QALYs or life expectancy at birth; examples include mental health disorders and learning disabilities. A focus on only one health outcome measure for all clinical areas will hinder the estimation of the opportunity cost for clinical areas where other factors are important. It is essential to use health outcomes that reflect the priorities of the health care system for each clinical area when estimating a threshold that reflects opportunity cost.

2.4. The structure of the paper

This paper further explores these three questions. We propose possible methods to address the issues raised by these questions. The paper is structured as follows:

- Section 3 explores the possible effect that inefficiencies at a health location have on the estimation of opportunity cost.
- Section 4 explores possible differences between health locations when they make decisions at the margin, and the effect of this on the threshold. This section also considers whether there exist elements suggesting variability across clinical areas.
- Section 5 looks at the importance of the selection of outcome measures.
- Section 6 comments on the role of stakeholders on the construction of an adequate methodology to measure the threshold.
- Section 7 concludes summarising the issues we hope our research will address.

3. ARE PURCHASERS ALWAYS DISPLACING SERVICES WHEN A NEW MEDICAL TECHNOLOGY IS ADOPTED? ROLE OF EFFICIENCY

3.1. Evidence concerning efficiency improvements

Karlsberg Schaffer et al. (2013) conducted a series of interviews with Scottish senior NHS finance managers to determine whether it was possible to identify specific services displaced (and therefore opportunity costs) when new technologies are introduced as a result of NICE guidance. They found that displacements in health services are rare. Increasing technical efficiency is the more common practice to achieve savings. This was confirmed in a second study conducted in Wales where the most common reported action to fund a new health treatment was to increase targets for efficiency savings. This implies savings achieved to fund a new treatment should not have a negative effect on other health and related benefits (Karlsberg Schaffer et al. 2016). This, of course, assumes that the technical efficiency gains can be made without impacting on any important service outcomes.

The potential for making efficiency gains means that a threshold estimate based on the assumption that each health location is displacing resources to fund new treatments will lead to the estimation of a threshold value that is lower than the real one. This is because the opportunity cost in terms of health outcomes of those health locations that can improve efficiency with low or zero impact on health outcomes will be low or zero in comparison to those that need to displace resources needed to satisfy demand for new health services[3]. This was previously noted by Barnsley et al. (2013) in their comment on the methodology of Claxton et al. (2013) to estimate the threshold value. Barnsley et al. (2013) suggest that if the assumption that health locations are operating efficiently is relaxed, then decreasing a budget over time may have little impact on health because the providers could provide the same services (and health outcomes) more efficiently.

[3] The threshold value is incremental cost / incremental outcome. If there is no displaced treatment, then the change in outcome is zero and the threshold value is infinite. Arguably the relevant threshold then becomes the value of what else could be done with the resources spent on the new technology. However, that is not the question being asked in the opportunity cost threshold debate.

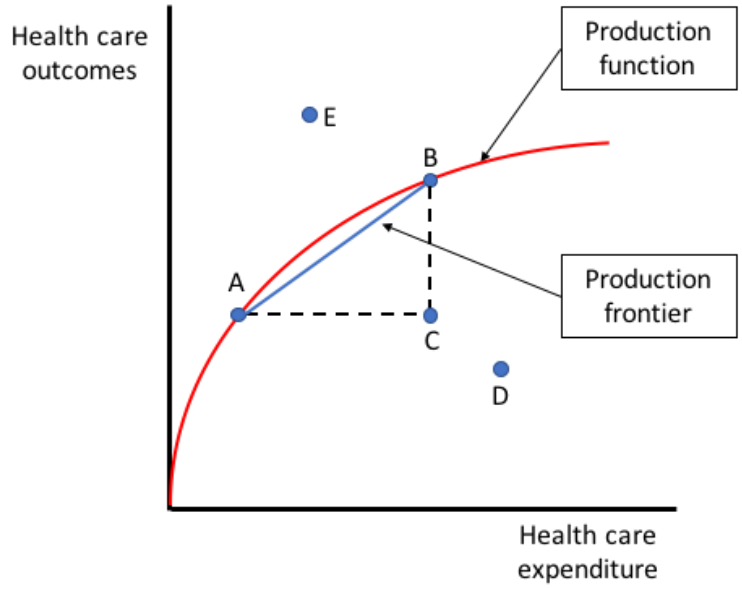
Studies of NHS efficiency suggest that there are differences in the productivity of health locations. For instance, Bojke et al. (2013) analysed health locations in England in the year 2007/8 and found that productivity varied from 5% above, to 6% below, the national average.

3.2. Using Data Envelopment Analysis to consider efficiency

As a result, the methodology to estimate the threshold based on the opportunity cost of adopting new health treatments should consider the possibility of current inefficiency in service provision, and that increased pressure on funds is associated with reduced inefficiency. Data Envelopment Analysis (DEA) is a potentially suitable approach, as it is possible to estimate the relationship between outcomes and expenditures while considering the efficiency / inefficiency of the service providers. DEA is a useful and robust methodology that has been employed in a large number of studies to understand the efficiency of health systems (for example Hollingsworth 2008; Tsai and Mar Molinero 2002; Cylus, Papanicolas, and Smith 2015; Pelone et al. 2014; Emrouznejad, Parker, and Tavares 2008). DEA allows the estimation of the effect of a marginal change in expenditures on the mix of health outcomes given different levels of efficiency of local health commissioners and providers. An important advantage of DEA is that it allows the consideration of multiple measures of outcome.

For simplicity, the concept of the DEA can be explained using an illustration based on four economics concepts: the *production function*, *technical efficiency*, *economic efficiency* and the *production frontier*. A **production function** shows the maximum outcome combination that can be produced during a specific time period given fixed resource inputs and a fixed production technology.

Figure 1. Economic efficiency



Technical efficiency is achieved when a production unit achieves this maximum output for a given level of physical resource inputs; it is not possible for the unit to produce more output without using more inputs, or to use fewer inputs without producing less output. If the volume of resource inputs is aggregated using a common value base,

their unit cost, then it is possible to construct a production function having a single input, as in the red line in Figure 1.

In this case, the production function shows not only technically efficient production, in terms of physical resources, but also **economic efficiency**, meaning the lowest possible cost of producing a given output level, or the greatest possible output for a given budget.

Units producing at points A and B in Figure 1 are efficient, but those at points C and D are not; for point C, for example, it would be possible to produce greater health outcomes with the same expenditure, as in point B, or have lower expenditure for the same outcomes, as in point A. Any point on the function between A and B would represent an unambiguous improvement, in terms of efficiency, for a unit at Point C. Point E, and any other point above the production function, represents a level of production that cannot be achieved with the current production technology.

Although units on points C and D are both inefficient in absolute terms, they differ in how inefficient they are. A unit at point D is less efficient than a unit at point C, since it has higher expenditure and lower output. A key aim of DEA is to estimate relative efficiency, measured by the distance that an inefficient unit is from the technically efficient production function. The distance between a line and a point is not unambiguously defined, so DEA uses a *Farrell radial measure*, which restricts the comparison to a point on the function that shrinks outputs and inputs in the same proportion as in the unit's real level of these.

Production functions represent production points that may not be possible to observe in the real world – for example there may be no production units that have the production levels that could theoretically be achieved. Instead, DEA and other estimation methods are based on the concept of a **production frontier**, also known as a *best-practice* frontier. Within a group of production units, this consists of all units that are not inefficient relative to any other unit. In Figure 1, ignoring the impossible point E, producers at points A and B are both more efficient than those at points C and D, but not inefficient relative to each other. They therefore form the production frontier, shown as a blue line, from which the relative inefficiency of the producers represented by points C and D can be measured.

Because DEA compares each producer, in our case English healthcare areas at Primary Care Trust (PCT) level with only the best producers, known as *peers*, problems related to analyses based on averages are avoided. DEA does not assume any functional form for the production function. An advantage of DEA is that allows analysis of both multiple inputs, and multiple outcomes, which is important in the analysis of health systems since it allows more than one health outcome indicator to be included, and therefore captures the complexity of each PBC and the effect of a change in health expenditures on the population.

The opportunity cost of an effective new health care technology can be regarded as a reduction in the budget available for other technologies, offset by an improvement in health outcomes. For producers A and B, a new technology offering the same cost-effectiveness ratio as the currently most marginal technology would enable them to substitute the technologies without impacting on overall cost-effectiveness. A new technology that has a better cost-effectiveness ratio than the currently most marginal would enable them to generate more output from the current budget. However, producers C and D could adopt the new technology even if it has a worse cost-

effectiveness ratio than current technologies if they are able to improve efficiency. Consequently, the opportunity cost of the entire system, understanding this as the sum of all opportunity costs, is higher when all PCTs are efficient than when there exist inefficient PCTs. An estimate of the opportunity cost of the currently most marginal technologies for the system as a whole, based on an average of both efficient and inefficient producers, will be higher than the true value.

Let us assume that the production function in Figure 1 has been derived using DEA to estimate a production frontier from a population of PCTs. Consider now a shock that results in a decrease in health expenditures by the distance AC. We consider the two PCTs: B and C. Observe that B's only option is a decrease in health outcomes. This is represented by a movement along the production function in Figure 1 from B to A. The PCT C has the option, however, of improving efficiency and decreasing inputs without affecting health outcomes. It could move to point A after reducing expenditures, a horizontal movement along the blue dotted line in Figure 1. Therefore, for some of the PCTs the opportunity cost of a decrease in health expenditures could, in theory, be zero if they can increase efficiency (as for C), but greater than zero if the possibilities of improving efficiency are exhausted (as for B). Consequently, the opportunity cost for the entire system of PCTs providing a new treatment, that is the sum of all opportunity costs, is higher when all PCTs are efficient than when there exist inefficient PCTs who increase efficiency in response to the cut in expenditure.

The sensitivity of DEA to outliers and influential DMUs has been extensively reported in the literature and more robust estimations methods have been proposed (e.g. Simar and Wilson, 2015). A further complication arises from the fact that health expenditures are affected by the quality of the health outcomes targeted by CCGs. This endogeneity introduces a new layer of interdependency between inputs and outputs which can be considered in DEA models (Mayston 2017).

Alternative methods have been proposed to compare efficiency across providers in the context of multiple outcomes. For example, Gutacker and Street 2017 estimate multivariate multilevel models to study hospital performance in terms of length of stay, readmission rates, post-operative patient-reported health status and waiting time.

It is important to be aware that, although DEA analysis is a useful tool to help estimate the opportunity cost of a health location, it is based on a number of simplifying assumptions, some of which could be considered limitations for the correct estimation of the opportunity cost. For instance, DEA is a deterministic rather than a statistical technique, therefore, results are sensitive to measurement error. In other word, when inputs from a particular health location are underestimated or overestimated, this health location could become an outlier. Incorrect outliers could significantly distort the shape of the production frontier and artificially reduce the efficiency scores of other health locations. It is important to consider how to treat potential outliers in the analysis (Gholam Abri, 2017).

In addition, efficiency scores estimated throughout the DEA are sensitive to sample size. This because health locations that do not have peers are considered efficient by default. When the sample size increases it is easy for the DEA algorithm to find comparable peers for health locations that were previously located in the frontier, which tends to reduce the average efficiency score (Pedraja-Chaparro et al., 1999). Similarly, when the number of health locations is constant but the number of outputs and/or inputs decreases, efficiency scores tend to decrease on average (Pedraja-Chaparro et al., 1999). This

responds to the fact that the number of dimensions in which a particular health location can be unique decreases. In order to estimate the opportunity costs of decreasing expenditures using the DEA, health expenditures should be included as unique inputs. This could affect the identification of variations in the use of particular inputs which could affect the efficiency scores. A possible solution to avoid this problem is to consider that health outcomes are the product of more than one year of expenditures by including more than one year of health expenditures.

Selecting the appropriate input and output variables to include in the model is the most critical decision to obtain reliable efficiency scores. In this regard, the literature suggest that the wrong selection of the variables could affect the DEA estimations (Smith, 1997; Simar and Wilson, 2001). For instance, the inclusion of irrelevant variables in the model, the omission of relevant inputs (Galagedera and Silvapulle, 2003; Ruggiero, 2005) or the presence of correlation between inputs could affect the results (Smith, 1997). In the case of the health system, lack of information in the health system could lead to the omission of important health outcomes. According with Smith (1997) high correlation between inputs allows DEA to yield better estimates of efficiency, however, the reductions in efficiency from omitted an input are higher than those arising when there is positive correlation between inputs. Therefore, a key element in the estimation of the DEA is to deeply consider the variables that should be included and test the correlation among those variables.

DEA approximates the production frontier by a piece-wise linear envelopment. In other words, DEA constructs a subset of the inputs and output space that contains all observed health locations. The piece-wise linear envelopment approximation of the production function is not smooth, continuous or differentiable. In the DEA model there will be corner points or points on edges and ridges where differentiability is not valid (Førsundet al. 2007). This increases the difficulty of estimating the outcome elasticities, and so, the opportunity cost.

The production assumptions of DEA include monotonicity (i.e. free disposability of inputs and outputs), convexity, and various notions of returns to scale. The monotonicity assumption implies that the output will be constant or increase, but never decrease when inputs increases. This is not always the case in healthcare. For instance, an increase in expenditures to finance a new screening program could lead to the detection of additional cases, which consequently could have a negatively effect on some health outcomes.

Thus, DEA is a useful tool but needs to be used in a way that minimises the risk of obtaining any misleading results.

4. WHAT SERVICES ARE DISPLACED AND WHERE? VARIATION IN OPPORTUNITY COST PER LOCAL HEALTH AUTHORITY AND CLINICAL AREA

4.1. Moving away from averaging across clinical areas and locations

In general, approaches to estimating the threshold have sought to estimate the average opportunity cost across all health locations. This approach overlooks potential differences between health locations, for example in the way in which health services are provided in different locations and the importance of each clinical area in terms of both number of

patients and health care provision. If there are such differences then health locations could be producing health gain in different ways, i.e. they will have different 'production functions', and will produce different health gains per pound spent. In other words, health locations that have dissimilar 'production functions' are likely also to have different opportunity costs. In addition, because of this, and because of different preferences, they will adjust different services when faced with the need to cut expenditure. The question then arises as to whether an averaging process helps or hinders our ability to understand what is displaced if a new treatment is adopted.

The results presented by Claxton et al. (2013) indicate differences in opportunity cost between health locations in England. They found that even after considering differences in input prices and needs, the cost of a life year in any given clinical area is lower in 'high per-capita spend' Primary Care Trusts (PCT) (those whose spend is greater than the average predicted spend in a particular PBC) than it is in 'low per-capita spend' PCTs (those whose spend in a particular PBC is less than the average predicted spend). They suggest that comparatively higher expenditures in particular PBCs can therefore be explained by relative cheaper additional units of health gains in a particular clinical area, in other words more is spent because efficiency is greater so that outcomes can be bought relatively cheaply. As they state, this implies that some PCTs are less efficient than others. Indeed, if everyone was on the frontier of the same production function in a particular PBC, we would expect diminishing marginal returns to additional spending in a clinical area. Claxton et al. (2013) found the opposite, implying that PCTs do not all have the same production functions.

Claxton et al. (2013) find that lack of available data made it impossible to jointly estimate opportunity costs for each clinical area in relation to the maximization of the total well-being per health location (allocating resources in a way that the last pound spent has equal marginal health related benefits in all clinical areas). The alternative was to separate the estimation of opportunity cost by clinical area. Claxton et al. (2013) found that opportunity cost differs considerably depending on the PBC / clinical area. For example, they estimated that the cost per life year gained for the PBC 'Genito-urinary problems' is over £850k whilst for 'Circulatory Problems' it was around £11k. The estimated threshold across PBCs varies from £2k per QALY in Respiratory disease to £2.9m per QALY in Maternity and Neonates. Table 1 sets out estimated PBC marginal cost per QALY in 2008/9. This suggests that if we consider each clinical area in isolation, there are important differences in opportunity cost.

We can put together the two indicators of heterogeneity: (i) differences between PCTs with PBC 'per-capita high spend' and PCTs with PBC 'per-capita low spend' production functions suggesting dissimilarities of opportunity cost as between health locations and (ii) differences in the ability to generate QALYs by clinical area. As a consequence, PCTs required to reduce expenditure may choose to do so in those PBC areas where they have to incur high costs to achieve health outcomes, so that they achieve the expenditure savings by reducing outcomes by less than if they spread the cuts across all PBC areas. This may reflect both the "national" differences in the ability to generate health gain per £ spent in a particular clinical area and local differences in health production functions.

Table 1. Estimated PBC marginal cost per QALY in 2008/9

Programme Budget Category	Cost per QALY	Programme Budget Category	Cost per QALY
Cancer	£16,997	Mental Health	£49,835
Circulatory	£7,038	Learning Disability	£78,854
Respiratory	£1,998	Problems of Vision	£76,850
Gastro-intestinal	£7,293	Problems of Hearing	£19,070
Infectious Diseases	£20,829	Dental Problems	£55,916
Endocrine	£3,124	Skin	£174,775
Neurological	£5,480	Musculoskeletal	£20,254
Genito-urinary	£43,813	Poisoning and AE	£163,766
Trauma & Injuries	NA	Healthy Individuals	£1,483,012
Maternity & Neonates	£2,969,208	Social Care Needs	NA
Disorders of Blood	£28,305	Other	NA

Source: Claxton et al. (2013), Table 5.2

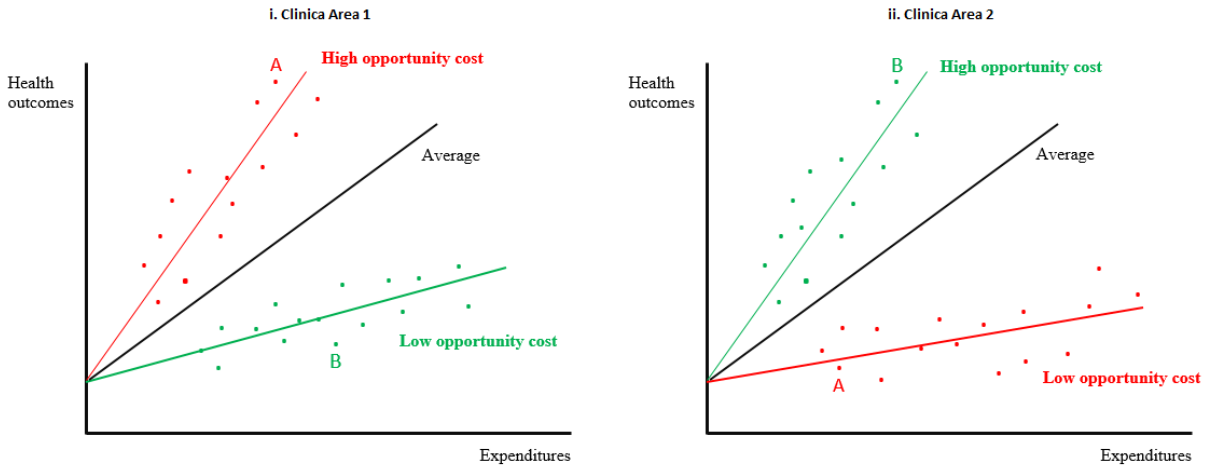
Figure 2 illustrates these ideas for Clinical Area 1 and Clinical Area 2. The slope of the black line represents a relationship between expenditure and outcomes estimated under the assumption that all health locations behave similarly. We can see than, on average, for simplicity, the two clinical areas have the same relationship between expenditure and health outcomes. However, in

Figure 2 we have two distinct groups of health locations: the red and the green group, with differing high and low opportunity costs in each clinical area.

In Clinical Area 1, the red group is currently more capable of increasing health outcomes for every pound they spend in comparison with the green health locations. They translate additional expenditure into improved outcomes at a much more efficient rate than the green health locations. However, in Clinical Area 2, the reverse is true. The red group is less able to increase health outcomes than the green group.

In Clinical Area 1 the red group has a higher outcome elasticity than the green group (the red line is steeper than the green line), meaning that the opportunity cost for the red group is higher than for the green group.

Figure 2. Hypothetical relationship outcomes and expenditures for different clinical areas



If the health locations are acting rationally and taking decisions at the margin by displacing the least cost-effectiveness technology, we will observe that a health location in the red group like *A* decreases health expenditures that affect Clinical Area 2 more than Clinical Area 1. This is because the negative effect on health outcomes will be lower in Clinical Area 2. The opposite will be true for a health location like *B* where most of the adjustment could be observed in Clinical Area 1. Consequently, the opportunity cost estimated on the basis of the slope of the black line may differ from the actual opportunity cost of diverting health resources from Clinical Area 1 and/or 2 to an alternative use. It would be more representative to have a value closer to that of the lower opportunity cost group than that of the average.

4.2. Evidence of variation between health locations

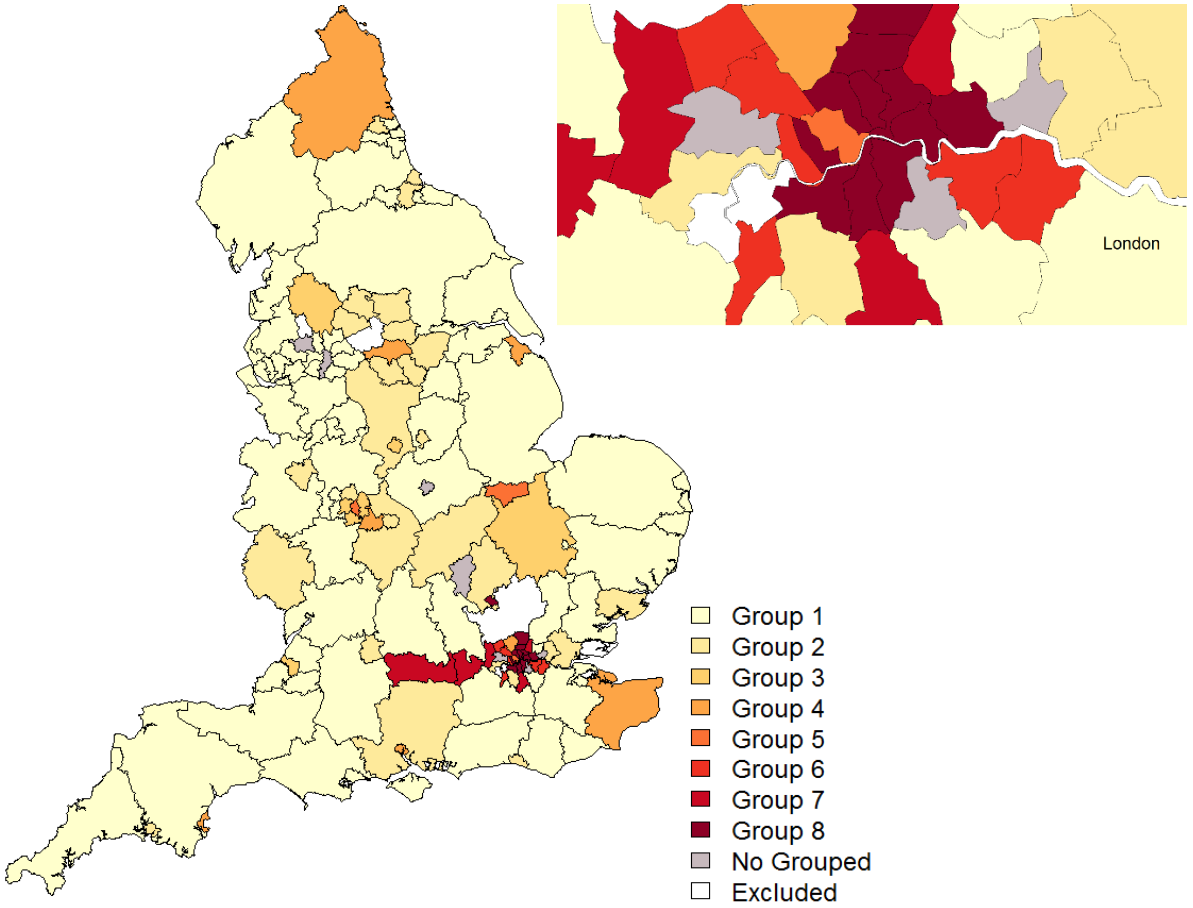
Divergence between the slope of the average of all health locations and the actual opportunity cost will only occur if there are significant differences in the outcome elasticities as between health locations and that these can lead to different decisions by clinical area. There is indeed evidence that health locations do behave differently. We include, in the Annex to this paper, results based on a cluster analysis, which found that there are important dissimilarities in expenditure among health locations in England, after adjusting for differences in needs and for unavoidable geographical cost variations.

Cluster analysis methodology allows the separation of observations into different groups reflecting the selected variables. This means that observation values for the selected variables within each group share a higher similarity than those of the members of any other group. In the analysis reported in the Annex, English health locations (defined by PCTs) were divided into different groups according to their expenditure patterns. These expenditure patterns were observed through the amount that each PCT spends in each one of the clinical areas (PBCs).

The observations are allocated to different clusters, such that cluster's members share a similar model of behaviour regarding the level of expenditures in each PBC. Based on four years of data (2008/08 - 2011/12) the results of the analysis indicated that health locations cannot in general be considered to behave similarly to one another. There is more than one cluster in each year, and these different clusters are stable over time. Eight groups of PCTs that share the same cluster during the four years were identified (Figure 3). This indicates that groups of English health locations are demonstrably different from one another. Within the clusters they are similar, but across the clusters they are quite different, suggesting that different PCTs are likely to have different production functions.

This finding means that it may not be appropriate to assume that the average effect can be taken as being equal to the marginal effect, i.e. that for a given clinical area a marginal decrease in health outcomes as a response to a marginal change in the expenditures is the same regardless of the PCT in which the expenditure reduction takes place.

Figure 3. Cluster analysis differences among local health location in England



Source: Authors' elaboration based data from the Department of Health (2008/09-2009/10-2010/11-2011/12)

4.3. Using a quantile regression function

Current approaches estimate the relationship between expenditures and health outcomes using a linear regression methodology (Claxton et al. 2013; Lichtenberg 2004; Martin, Rice, and Smith 2008, 2012). This is a method for estimating the expected mean of a dependent variable (in our case health outcomes) conditional on the levels or values of independent variables (e.g. health expenditures). However, if the decisions made as a result of the approval of a new health technology are considered at the margin and not at the mean and the health locations are heterogeneous, a methodology that does not focus on the mean of the health outcome is preferable.

One possible approach is to employ a quantile regression function which can accommodate non-linearity in the relationship between expenditures and outcomes, and variation in outcome elasticities at the margin. This can be done by estimating models for several conditional quantile functions (Koenker 2005). Quantile regression can be seen as an addition to the Ordinary Least Square (OLS) model. OLS is based on the estimation of a conditional mean model whereas quantile regression is based on an estimation of a group of models for several conditional quantile functions. Quantile functions are estimated by minimizing an asymmetrically weighted sum of absolute errors. In the situation in which health locations act differently, this approach offers the ability to identify the low and high opportunity cost curves presented in

Figure 2, which otherwise would be hidden in an average. Previous studies have examined the relationship between health expenditures and health outcomes using quantile regression (Lê Cook and Manning 2013; Obrizan and Wehby 2012; Piérard 2016). Obrizan and Wehby (2012) found significant heterogeneity in the effect of expenditures on life expectancy. This heterogeneity underestimates (overestimates) the expenditure returns for locations ranking in low (high) life expectancy quantiles. A classic OLS approach would not be able to capture any of these differences.

Alternatively, health production functions have been estimated using panel data techniques to identify productivity dispersion across health locations on the basis of a single outcome, risk-adjusted survival, as well as the relationship between the degree of adoption of technological diffusion in health locations and marginal productivity (Skinner and Staiger 2015), or the relationship between market allocation and hospital performance (Chandra et al. 2016).

As argued above and illustrated in Figure 2, DEA can estimate the opportunity cost value for each one of the health locations. DEA analysis has the advantage of comparing each DMU (in our case health locations) with only the most efficient DMUs, which avoids problems resulting from analyses based on averages. Moreover, DEA does not assume any particular process in which the inputs must be combined to produce multiple outputs (there is not a functional form for the production function). Dowd et al. (2014) suggest that DEA input efficiency might provide a useful scalar measure of value associated with higher quality and lower costs which may serve to inform value-based payments.

5. USING HEALTH OUTCOMES REFLECTING THE HEALTH SYSTEM PRIORITIES

Maybe the most important assumption in the estimation of the threshold is the selection of the relevant set of health outcomes. An estimate of the threshold should correctly reflect the opportunity cost for the health system of the inclusion of a new health treatment. This depends on the prioritisation of outcomes according to policy makers' and society's preferences. For instance, if healthcare access improvements are given a higher priority (that is, have greater social value) than emergency services improvements, displacements that negatively affect healthcare access will be perceived as having a higher opportunity cost than displacements that affect emergency services. A threshold that is estimated without taking into account the full set of outcomes that are of value will not reflect the true effect on outcomes of increasing or decreasing health expenditures. Therefore, it is important to include in the analysis *all* of the outcomes that are desired from health care expenditures.

Common practice is to quantify the threshold in terms of the cost per QALY gained, as in the majority of health economic evaluations used to inform NICE decision making. Indeed, this is the approach specified by NICE. However, there is a lack of evidence that allows the estimation of the QALYs produced by health location and by clinical area. This hinders the estimation of a threshold value using QALYs. Claxton et al. (2013) seek to overcome this problem by assuming that health locations are as good at improving morbidity (quality of life) as they are at reducing mortality. It is easier to obtain information on mortality rates than morbidity. They were able to estimate the opportunity cost in terms of mortality and then transform it to QALY values by making the assumption that the effect of health expenditure on morbidity was proportional to its effect on mortality, with the proportions coming from estimates of the mortality and

morbidity components of the burden of disease in the relevant clinical area. This work-around solution has limitations. Barnsley et al. (2013) highlight the difficulties of moving from estimates of mortality to impute additional quality of life gains to arrive at QALYs gained. They argued that there is no evidence to support the use of a 1:1 proportional relationship between health location capacity for reducing mortality rates and its capacity for improving morbidity.

In similar work undertaken by Vallejo-Torres, García-Lorenzo, and Serrano-Aguilar (2016), the impacts of health expenditure on mortality and on health related quality of life (HRQoL) respectively are separately and directly estimated using a health outcome variable capturing differences in mortality and morbidity. This was possible because Vallejo-Torres, García-Lorenzo, and Serrano-Aguilar (2016) estimated only one relationship between expenditures and health outcomes, and used it as an average for all clinical areas and health locations. They used HRQoL information, which were available by Spanish region but not disaggregated by clinical area. Their approach suffers from the weakness that a threshold estimated based on the average ignores that decisions are made at the margin. However, they avoided the need to use mortality data to estimate HRQoL effects.

There is an implicit assumption in all of these analyses that health locations' priorities are well reflected in QALYs gains (Claxton et al. 2013; Martin, Rice, and Smith 2008, 2012; Vallejo-Torres, García-Lorenzo, and Serrano-Aguilar 2016), if only we could estimate them properly. We need to avoid confusing what we might think as health economists is the ideal normative assumption for our analysis with what health systems such as the NHS in England try in practice to achieve. An assumption of QALY maximisation ignores other evident priorities such as equity, and improving healthcare access by, for example, reducing waiting times (Shah et al., 2012). Health locations could be maximizing a completely different set of health outcomes depending on the clinical area. Some of the outcomes might be fully captured by QALYs and others not. Barnsley et al. (2013) highlighted that the large marginal cost per QALY differences among the PBCs estimated by Claxton et al. (2013) suggested that health locations in England are seeking to maximise a more complex set of outcomes than QALYs alone. Shah et al. (2012) found that the national Department of Health also has multiple objectives beyond QALY gain.

Table 2 shows a set of health outcome measures by clinical area which are currently available for the English health care system. These reflect some of the measures that decisions makers could be looking at when making decisions related to the displacement of services. Some of them, such as the effect on HRQoL of changes in the number of people suffering from asthma, could be readily translated into QALYs gained, whilst others refer to the numbers of people accessing health care services or other outcomes, such as increased employment of people with mental illness, that are measured in units that cannot easily be translated into QALY gains.

Table 2. Possible measures of health outcomes: Example of a subgroup of PBCs

Code	Main Category	Possible outcome
2	Cancers and Tumours	1) One-year survival from all cancers 2) One-year survival from breast, lung and colorectal cancers 3) Record of stage of cancer at diagnosis 4) Percentage of cancers detected at stage 1 and 2 5) Record of lung cancer stage at decision to treat 7) Palliative care
5	Mental Health Disorders	1) Depression (ages 18+) (number of people on the QOF register for this condition) 2) Unplanned readmissions to mental health services within 30 days of a mental health inpatient discharge in people aged 17 and over 3) Excess under 75 mortality rate in adults with serious mental illness 4) Mental Health (number of people on the QOF register for this condition) 5) Dementia (number of people on the QOF register for this condition) 6) Employment of people with mental illness
6	Problems of Learning Disability	1) Learning Disabilities (ages 18+)(number of people on the QOF register for this condition)
11	Problems of the respiratory system	1) Asthma (number of people on the QOF register for this condition) 2) Unplanned hospitalisations for asthma, diabetes and epilepsy in under 19s 3) Emergency admissions for children with lower respiratory tract infections 4) Chronic Obstructive Pulmonary Disease COPD (number of people on the QOF register for this condition) 5) The percentage of people with COPD and MRC Dyspnoea Scale >=3, identified on GP systems, referred to pulmonary rehabilitation
15	Problems of the Musculoskeletal system	1) Rheumatoid arthritis (16+) (number of people on the QOF register for this condition) 2) Osteoporosis (ages 50+) (number of people on the QOF register for this condition) 3) Elective knee replacement (Primary) procedures - PROMS 4) Elective Hip replacement (Primary) procedures - PROMS 5) Hip fracture: incidence

QOF - Quality Outcomes Framework. COPD - Chronic Obstructive Pulmonary Disease. MRC - Medical Research Council. PROMS - Patient Reported Outcomes Measures. Source: Outcomes from NHS England (2016) and classified according with Department of Health (DH) (2012)

The estimation of threshold values requires consideration of a broader set of health outcomes than QALYs to better reflect the reality of the health system. A characteristic of the DEA approach discussed earlier is that it not only allows the analysis of a multiple inputs process, but also of a multiple outputs process. This is an important advantage in the analysis of the health care system since it allows the consideration of more than one health outcome indicator.

6. STAKEHOLDERS' ROLES IN DETERMINING OPPORTUNITY COSTS

Opportunity cost is defined in different ways. For example "the opportunity cost of committing resources to produce a good or service is the benefits foregone from those same resources not being used in their next best alternative" (Morris, Devlin and Parkin 2007 p.3), and as "the cost of a benefit that must be forgone in order to pursue an alternative" (Becker, Ronen, and Sorter 1974). The assumption in regard to the threshold value is that the concept of opportunity cost is understood by policy makers in terms of costs and benefits for the health system. No improvement in methodologies to

estimate the threshold will help if decision makers do not consider costs and benefits in their decisions and hence will not use the estimate of the threshold.

How are resource allocation decisions made in practice, and to what extent does economic evaluation play a part? The link between actual decisions and the consideration of economic evaluation in the health system to make decisions has generated substantial discussion. Eddama and Coast (2009) concluded that economic evaluations are not considered at the local level and explained that this is a consequence of decisions being more related to management factors (e.g. employment of extra staff) than to the acquisition of medicines or specific interventions. Their study suggests that there is a general lack of awareness of the economic evaluation approach to decision-making and a misunderstanding on how economic evaluation can be useful locally. Similarly, Appleby et al. (2009) find that a large number of decisions made by local health authorities are related to service reconfiguration and demand management initiatives. However, they propose that a detailed analysis of costs and benefits of every health service is not required for the estimation of the threshold value to make sense in reality. They state that what is required is that policy makers act as if such deep analysis has been done. This means that policy makers should know the most sensitive clinical areas where health outcomes would be highly affected by a reduction in health expenditures and act accordingly. Their results indicate that health locations consider cost-effectiveness in the decision making process mainly to remove procedures that are not technically efficient and to adopt cheaper procedures that are at least as effective as the standard treatment, i.e. when one treatment dominates the other in terms of cost-effectiveness.

The decision making process is not the same across different stakeholders who may have different ideas in terms of priorities, relevant costs and relevant outcomes. This will be translated into differences in opportunity cost caused by the displacement of health resources, which adds a new level of difficulty to the threshold estimation.

Therefore, it is fundamental that any attempt to estimate the opportunity cost of a new medical technology to the health system incorporates the opinion and preferences of stakeholders. There is a need for understanding which outcomes are considered more or less important. The identification of the main actors in the decision making process, and which perspectives are shaping the agenda for the inclusion of new health treatment, is required to understand the nature of the opportunity cost. This is only possible by engaging these different stakeholders. This is preferably done during the construction of the methodology and identification of data collection needs, but is most important for the validation of the results.

7. CONCLUSIONS

Establishing appropriate empirical thresholds remains a challenge for all countries and health care systems. In order to contribute to the design of a methodology that could be adopted by policy makers during the decision making process, we explored some implicit and explicit assumptions that hinder the correct estimation of the threshold under the current methods being used.

First, we considered the role of possible health location inefficiencies on the estimation of the opportunity cost, and consequently, the threshold value. When inefficiency exists and is ignored, this could result in valuations of the threshold that understate the relevant level. There are indications that inefficiencies exist among health locations. However, further research is needed to estimate and understand these inefficiencies. It

could be the case that inefficient health locations will not be able in the short run to reduce their inefficiencies, or, alternatively, that any improvements that would result if they did, would be insufficient to release enough funds to cover a new health technology. In these circumstances, displacements could still happen. Therefore, the estimation of the threshold value should allow for observation of the actual level of inefficiencies as well as an ability to consider the previous capacity of health locations to respond to changes in expenditures.

Second, we considered the effect on the estimation of the opportunity cost if health locations decide on spending priorities in different ways. If such differences exist, disparate marginal displacements with different opportunity costs could be happening throughout the health system. Methodologies such as cluster analysis and quantile regression can shed light on the existence of significant heterogeneity among health locations which should be considered in preparing an estimate of the threshold.

Third, we discuss the importance of selecting the appropriate set of health outcomes, such that they reflect health system priorities. Otherwise, we would be estimating a threshold that does not reflect likely displacement.

DEA can help to address these three problems relating to the estimation of the opportunity cost, by considering efficiency, allowing the estimation of the opportunity cost per health location, and including more than one outcome measure. DEA methodology has not been tested in this context and the issues we raised earlier would need to be addressed. It can, however, produce estimates of the opportunity cost by clinical area. The results of this analysis can be discussed with relevant stakeholders, to understand the validity of the results.

REFERENCES

Appleby, John, Nancy Devlin, David Parkin, Martin Buxton, and Kalipso Chalkidou. 2009. 'Searching for cost effectiveness thresholds in the NHS', *Health Policy*, 91: 239-45.

Barnsley, Paul, Adrian Towse, Sarah Karlsberg Schaffer, and Jon Sussex. 2013. 'Critique of CHE research paper 81: methods for the estimation of the NICE cost effectiveness threshold', Occasional Paper 13/01. Office of Health Economics.

Becker, Selwyn W, Joshua Ronen, and George H Sorter. 1974. 'Opportunity costs-an experimental approach', *Journal of Accounting Research*: 317-29.

Bojke, Chris, Adriana Castelli, Andrew Street, Padraic Ward, and Mauro Laudicella. 2013. 'Regional variation in the productivity of the English National Health Service', *Health Economics*, 22: 194-211.

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson. 2016. 'Healthcare exceptionalism? Productivity and allocation in the US healthcare sector', *American Economic Review*, 106: 2110-44.

Claxton, K., S. Martin, M Soares, N. Rice, E. Spackman, S. Hinde, N. Devlin, Smith P.C, and M. Sculpher. 2013. "Methods for the estimation of the NICE cost effectiveness threshold: Revised report following referees comments." In. North Yorkshire: University of York.

Claxton, K., S. Martin, M Soares, N. Rice, E. Spackman, S. Hinde, N. Devlin, Smith P.C, and M. Sculpher. 2015. "Methods for the estimation of the National Institute for Health and Care Excellence cost effectiveness threshold." *Health Technology Assessment*. 2015 Feb;19(14):1-503, v-vi. doi: 10.3310/hta19140

Cleemput, Irina, Mattias Neyt, Nancy Thiry, Chris De Laet, and Mark Leys. 2011. 'Using threshold values for cost per quality-adjusted life-year gained in healthcare decisions', *International Journal of Technology Assessment in Health Care*, 27: 71-76.

Cooper, W.W., Seiford, L.M. and Zhu, J. eds. 2011. Handbook on data envelopment analysis (Vol. 164). Springer Science & Business Media.

Culyer, Anthony, Christopher McCabe, Andrew Briggs, Karl Claxton, Martin Buxton, Ron Akehurst, Mark Sculpher, and John Brazier. 2007. 'Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence', *Journal of Health Services Research & Policy*, 12: 56-58.

Cylus, Jonathan, Irene Papanicolas, and Peter C. Smith. 2015. 'Using data envelopment analysis to address the challenges of comparing health system efficiency', *Global Policy,* 8(52): 60-68.

Department of Health. 2008/09-2009/10-2010/11-2011/12. 'Programme budgeting tools and data'. http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Managingyourorganisation/Financeandplanning/Programmebudgeting/DH_075743.

Department of Health (DH). 2010. 'Exposition book 2009-10 and 2010-11', Department of Health (DH). http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Managingyourorganisation/Financeandplanning/Allocations/DH_091850.

Department of Health (DH). 2011. 'Exposition book 2011-2012', Department of Health (DH), Accessed 11/01. https://www.gov.uk/government/publications/exposition-book-2011-2012.

Department of Health (DH). 2012. 'Overview of the programme budgeting methodology - programme budgeting mappings and definitions'. https://www.gov.uk/government/publications/overview-of-the-programme-budgeting-methodology--2.

DH Financial Planning and Allocations Division. 2011. "Resource allocation: Weighted capitation formula (Seventh Edition)." In. London: Financial Planning and Allocations.

Dowd, Bryan, Tami Swenson, Robert Kane, Shriram Parashuram, and Robert Coulam. 2014. 'Can data envelopment analysis provide a scalar index of 'Value'?', *Health Economics*, 23: 1465-80.

Eddama, Oya, and Joanna Coast. 2009. 'Use of economic evaluation in local health care decision-making in England: a qualitative investigation', *Health Policy*, 89: 261-70.

Emrouznejad, Ali, Barnett R. Parker, and Gabriel Tavares. 2008. 'Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA', *Socio-Economic Planning Sciences*, 42: 151-57.

Gallet, Craig A, and Hristos Doucouliagos. 2017. 'The impact of healthcare spending on health outcomes: A meta-regression analysis', *Social Science & Medicine*, 179: 9-17.

George, Bethan, Anthony Harris, and Andrew Mitchell. 2001. 'Cost-effectiveness analysis and the consistency of decision Making', *PharmacoEconomics*, 19: 1103-09.

Gutacker, Nils, and Andrew Street. 2017. 'Multidimensional performance assessment of public sector organisations using dominance criteria', *Health Economics*, 27(2): e13-e27.

Gholam Abri, A.,(2017). Impact of Outliers in Data Envelopment Analysis. International Journal of Industrial Mathematics. 9(4), 319-332.

Hollingsworth, Bruce. 2008. 'The measurement of efficiency and productivity of health care delivery', *Health Economics*, 17: 1107-28.

Førsund, F.R., Hjalmarsson, L., Krivonozhko, V.E. and Utkin, O.B.. 2007. Calculation of scale elasticities in DEA models: direct and indirect approaches. J*ournal of Productivity Analysis*, 28(1-2): 45-56.

Karlsberg Schaffer, S., J. Sussex, N. Devlin, and A. Walker. 2013. "Searching for cost-effectiveness thresholds in NHS Scotland." Research Paper 13/07. Office of Health Economics.

Karlsberg Schaffer, S., J. Sussex, N. Devlin, and A. Walker. 2015. 'Local health care expenditure plans and their opportunity costs', *Health Policy*, 119 (9): 1237-1244.

Karlsberg Schaffer, Sarah, Jon Sussex, Dyfrig Hughes, and Nancy Devlin. 2016. 'Opportunity costs and local health service spending decisions: a qualitative study from Wales', *BMC Health Services Research*, 16: 103.

Koenker, Roger. 2005. *Quantile regression* (Cambridge university press).
Lê Cook, Benjamin, and Willard G Manning. 2013. 'Thinking beyond the mean: a practical guide for using quantile regression methods for health services research', *Shanghai archives of psychiatry*, 25: 55.

Lichtenberg, Frank R. 2004. 'Sources of U.S. longevity increase, 1960–2001', *The Quarterly Review of Economics and Finance*, 44: 369-89.

Martin, Stephen, Nigel Rice, and Peter C Smith. 2008. 'Does health care spending improve health outcomes? Evidence from English programme budgeting data', *Journal of Health Economics*, 27: 826-42.

Martin, Stephen, Nigel Rice, and Peter C Smith. 2012. 'Comparing costs and outcomes across programmes of health care', *Health Economics*, 21(3): 316-37.

Mayston, David J. 2017. 'Data envelopment analysis, endogeneity and the quality frontier for public services', *Annals of Operations Research*, 250: 185-203.

NHS England. 2016. 'CCG outcomes tools'. https://www.england.nhs.uk/resources/resources-for-ccgs/ccg-out-tool/.

O'Mahony, James F, and Diarmuid Coughlan. 2016. 'The Irish cost-effectiveness threshold: does it support rational rationing or might it lead to unintended harm to Ireland's health system?', *PharmacoEconomics*, 34: 5-11.

Obrizan, Maksym, and George L Wehby. 2012. "Health expenditures and life expectancy around the world: A quantile regression approach." Discussion Paper Series No.47. Kyiv School of Economics and Kyiv Economics Institute.

Pelone, Ferruccio, Dionne Sofia Kringos, Alessandro Romaniello, Monica Archibugi, Chiara Salsiri, and Walter Ricciardi. 2014. 'Primary care efficiency measurement using data envelopment analysis: A systematic review', *Journal of Medical Systems*, 39: 1-14.

Pedraja-Chaparro, F., J. Salinas-Jiménez and P. Smith, 1999. On the quality of the data envelopment analysis model. *Journal of the Operational Research Society*, 50(6): 636-644.

Piérard, Emmanuelle. 2016. 'The effect of health care expenditures on self-rated health status and the Health Utility Index: Evidence from Canada', *International Journal of Health Economics and Management*, 16: 1-21.

Raftery, James. 2009. 'Should NICE's threshold range for cost per QALY be raised? No', *BMJ*, 338: b185.

Ruggiero, J., 2005. Impact assessment of input omission on DEA. *International Journal of Technology and Decision Making*, 4(3): 359-368.

Schwarzer, Ruth, Ursula Rochau, Kim Saverno, Beate Jahn, Bernhard Bornschein, Nikolai Muehlberger, Magdalena Flatscher-Thoeni, Petra Schnell-Inderst, Gaby Sroczynski, and Martina Lackner. 2015. 'Systematic overview of cost–effectiveness thresholds in ten countries across four continents', *Journal of Comparative Effectiveness Research*, 4: 485-504.

Simar, L. and Wilson, P.W.. 2001. Testing restrictions in nonparametric efficiency models. C*ommunications in Statistics-Simulation and Computation*, 30(1):159-184.

Smith, P.. 1997. Model misspecification in data envelopment analysis. A*nnals of Operations Research*, 73: 233-252.

Skinner, Jonathan, and Douglas Staiger. 2015. 'Technology diffusion and productivity growth in health care', *Review of Economics and Statistics*, 97: 951-64.

Sussex, Jon, and Karla Hernandez-Villafuerte. 2015. "Model of behaviour within fuzzy budget constraints." Office of Health Economics.

Tsai, P. F., and C. Mar Molinero. 2002. 'A variable returns to scale data envelopment analysis model for the joint determination of efficiencies with an example of the UK health service', *European Journal of Operational Research*, 141: 21-38.

Vallejo-Torres, Laura, Borja García-Lorenzo, Iván Castilla, Cristina Valcárcel-Nazco, Lidia García-Pérez, Renata Linertová, Elena Polentinos-Castro, and Pedro Serrano-Aguilar. 2016. 'On the estimation of the cost-effectiveness threshold: Why, what, how?', *Value in Health*, 19: 558-66.

Vallejo-Torres, Laura, Borja García-Lorenzo, and Pedro Serrano-Aguilar. 2016. "Estimating a cost-effectiveness threshold for the Spanish NHS." In *Spanish Health Economics Study Group Meeting, Gran Canaria*.

Weinstein , Milton C., and William B. Stason 1977. 'Foundations of cost-effectiveness analysis for health and medical practices', *New England Journal of Medicine*, 296: 716-21.

Weinstein, Milton, and Richard Zeckhauser. 1973. 'Critical ratios and efficient allocation', *Journal of Public Economics*, 2: 147-57.

Zhang, Ning, Li-Jun Ji, and Ye Li. 2017. 'Cultural differences in opportunity cost consideration', *Frontiers in psychology*, 8: 45.

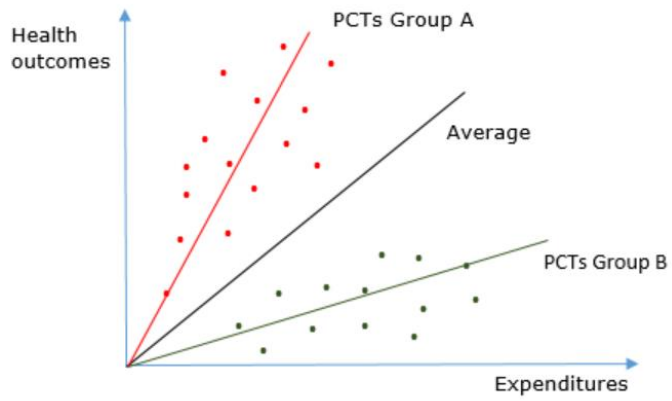
ANNEX 1: CROSS SECTION ANALYSIS OF MARGINAL NHS EXPENDITURE BY PCT AND PBC

A.1.1. Introduction

Different English health locations spend very different amounts per clinical area (also known as Programme Budget Categories (PBCs)) even after adjusting for population size (Claxton et al. 2013). Estimations to date of the effect on health outcomes of a change in expenditures on the NHS in England assume that all Primary Care Trusts (PCTs – being the local health authorities responsible at the time of the study for allocating NHS resources to different PBCs) behave in roughly the same way as each other when faced with a cut in available funds (Claxton et al. 2013; Martin, Rice, and Smith 2012). This overlooks potential differences in the way in which health services are provided in different localities and the relative importance of each PBC to a particular PCT in terms of the relative number of patients and relative size of health care provision. If there are such differences then different PCTs will be producing health gain in different ways (they will have different 'production functions' in economic language), and will produce different health gains per pound spent. When a PCT's available funds are reduced, its adjustment behaviour will reflect its healthcare priorities and its particular 'production function', i.e. the effect on the health outcomes in its geographical area of changing expenditure on a specific clinical service. In this case, the assumption that, for a given PBC, a marginal decrease in health outcomes as a response to a marginal change in the expenditures will be the same regardless of the PCT is not plausible. Empirical studies of the NHS in different parts of the UK suggest that different Health Boards, and hence probably PCTs, do not react to the same funding challenges in the same way (Karlsberg Schaffer et al. 2015; Karlsberg Schaffer et al. 2016).

If PCTs have different priorities, they will probably behave differently. However, it is not necessarily the case that PCT all have different healthcare priorities. Many PCTs are likely to have similar priorities. However, they may have different production functions. In this situation the average effect on health outcomes that results from a decrease in expenditure would not represent any particular part of the health sector. As an example, Figure A.1 shows a hypothetical case in which the PCTs can be divided into two different groups, one in which changes in expenditures considerably affect outcomes (PCTs Group A) and the other in which the marginal effect on outcomes is small (PCTs Group B). The average, represented by the black line, does not reflect what is happening in the healthcare system.

Figure A.1. Different groups of PCTs



We undertook an exploratory analysis to test the assumptions that PCTs are similar to each another in the decisions that they make. To determine whether the then 151 PCTs in England fall into distinct groups rather than all being members of, effectively, one large group, we applied Cluster Analysis Methodology. This looks at the degree of heterogeneity as between PCTs and then – depending on what is found - divides PCTs into different clusters on purely statistical grounds, such that the characteristics of the PCTs within each cluster are similar, but the clusters are dissimilar to each other. The characteristics we look for are whether the PCTs within one cluster are similar to each other in terms of the expenditures (adjusted by cost differences and needs) allocated to each of the 23 PBCs but different from other clusters. Four different sets of clusters were estimated, one set based on the 2008/09 data, one on the 2009/10 data, one on 2010/11 data and one on the 2011/12 data. By doing this, we were able to compare the composition of the different clusters over time. A first strong indication that PCTs cannot be considered to behave similarly to one another would be the existence of (i) more than one cluster in each year and (ii) stability of the PCT membership of clusters over time.

A.1.2. Data

The data base comprises the expenditures on own populations that each one of the 151 PCTs has made in each of the 23 Programme Budget Categories. The information comes from the "Programme Budgeting reference cost based PCT benchmarking workbook" of the Department of Health (Department of Health (DH) 2012). This data base is available over a number of years. Our analysis focuses on four financial years: 2008/09, 2009/10, 2010/11 and 2011/12. Between 2008-2012 the NHS in England was divided geographically into 151 PCTs, which in turn were grouped into 10 Strategic Health Authorities (SHAs). The PCTs were responsible for spending around 80% of the NHS England budget. Since the beginning of 2013 this structure was replaced with the creation of 213 Clinical Commissioning Groups (CCGs). Given the difference in the composition and percentage of budget handled between the PCTs and the CCGs the analysis presented in this Annex focuses only on the patterns of expenditures by PCTs.

Given that the objective is to analyse patterns of expenditures to observe possible differences in the health production processes, it is essential to consider the unavoidable dissimilarities between the PCTs that result in unavoidable differences in expenditures. In order to do that, the population of each PCT has been adjusted in three ways before the estimation of the per-capita expenditure:

- Firstly, the expenditure data in each year was adjusted using the Market Forces Factor (MFF) indices estimated by the DH in the "Resource Allocation Weighted Capitation Formula" (DH Financial Planning and Allocations Division 2011). This is because there are significant differences in prices throughout the UK. An analysis that does not take this into consideration will not reflect volumes of service expenditure, but the fact that some PCTs spend more because they face higher input prices.
- Secondly, the expenditure data were adjusted by the Emergency Ambulance Cost Adjustment (EACA) index, which considers the unavoidable cost variations of delivering emergency ambulance services (DH Financial Planning and Allocations Division 2011).
- Thirdly, it is necessary to consider the differences in health care demand that each PCT faces because of the underlying health of its population. There is not the same level of financial pressure in a location with a high percentage of elderly people when

compared with a location in which the majority of the population are less than 30 years old. In view of this, we make an adjustment using the methodology applied by the Department of Health in the PCT recurrent revenue allocations exposition book (Department of Health (DH) 2010, 2011; DH Financial Planning and Allocations Division 2011). The adjustment based on this methodology takes into consideration: acute needs; gender and age distributions; maternity needs; mental health needs; health inequalities; and HIV/AIDS prevention, treatment and care.

Finally, as our main interest is to observe whether PCTs can be grouped according to their `production functions´, we need in doing so to look at how PCTs are allocating their budgets as between clinical areas. Absolute per capita expenditure reflects both the importance of each clinical area for a PCT, but also the population needs faced by that PCT. A grouping based on this measure will probably cluster PCTs with comparable needs together, hindering the possibility of observing similarities and differences in the resources allocated as between clinical areas. Therefore, we use the percentage allocation of the total per capita expenditure as between PBCs rather than the absolute value of the per capita expenditures on each PBC. These percentages provide a clearer picture of how the total budget assigned to a particular PCT is distributed among the different clinical areas.

A.1.3. Methods and characteristics of the model to be estimated

Cluster analysis methodology allows the separation of the observations into different groups using selected variables. This means that observations of the variables within each group share a higher similarity in terms of the values of the selected variables, as compared with the members of any other group. In the case of this cluster analysis, the objective is to split the PCTs into different groups according to their expenditures patterns. These expenditure patterns are observed through the percentage of the total per-capita expenditure that each PCT spends in each one of the PBCs. After dividing the observations into different clusters, we can surmise that cluster members will share a similar model of behaviour regarding the level of expenditures in each clinical area (PBC).

Three PBCs are excluded from the analysis `Healthy Individuals´ (PBC number 21), `Social Care Needs´ (PBC number 22) and `Other´ (PBC number 23). This is because expert opinion suggests that the procedures for registering expenditures related to these three PBCs varies considerably during the period of interest.

To be able to cluster the observations into similar groups, it is necessary to define a unique comparable value that reflects differences between each observation in terms of the structure of the selected variables. In this case, we are looking at a value that reflects differences in expenditure patterns. The literature suggests that an estimate of what is termed the `Euclidean distance´ can be used in cases when the selected group of variables corresponds to a continuous measurements of a roughly linear scale. The PCT expenditures in each of the PBCs have this characteristic, so we can use as estimate of the "Euclidean distance." This is estimated as:

$$d_{ij}^{t} = \left[\sum_{k=1}^{n}\left(x_{ik}^{t} - x_{jk}^{t}\right)^{2}\right]^{1/2} \qquad (1)$$

where the x_{ik}^{t} and x_{jk}^{t} are the expenditures in *kth* PBC (k = 1,…,23) undertaken by the PCTs *j* (j = 1,…,151) and *i* (j = 1,…,151) during the year *t*.

Lower values of the distance d_{ij}^{t} means that PCTs *j* and *i* are similar in terms of patterns of expenditure. We are also interested in the study of the cluster stability over the four year of analysis. Therefore, for each pair of PCTs four distances are estimated:

d_{ij}^{08-09}, d_{ij}^{09-10}, d_{ij}^{10-11} and d_{ij}^{11-12}.

There are a number of algorithms to construct clusters based on the distance. In order to obtain robust results, we decided to apply three different algorithms and to test the stability and proximity of the clusters. The algorithms applied are the k-means, k-medoids, and the hierarchical method. We test the robustness and validity of the grouping resulted from each of the three algorithms through two indicators. First, the Dunn Index (DI), that measures the proximity of the observations inside the cluster (looking for small variances between members of the cluster) and the distance with the members of other clusters (expecting to see that the means of different clusters are sufficiently apart). Second, the SD Validity Index, which tests the average compactness or scattering of the clusters (i.e. the variance within the clusters) and the degree of total separation between the clusters.

An important part of the cluster analysis is related to the decision as to the number of clusters in which the observations are grouped. In our case, the optimal number of clusters has been approximated by estimating the "silhouette plot" which estimates how close each point in one cluster is as compared to points in the neighbouring clusters. The optimal number of clusters is estimated based on achieving the optimum average silhouette width. In addition, we also use a "dendrogram plot", which is a representation of appropriate hierarchical clustering, to explore the potential value of a hierarchical approach.

During the analysis of the optimal number of clusters, four PCTs that behave extremely differently to the rest of the country were identified: 1) Richmond and Twickenham, 2) Kirklees, 3) South East Essex and 4) Hertfordshire PCT. These observations are considered outliers and excluded from the database since this could affect the construction of the clusters.

After the exclusion of the four outliers, the possible optimal number of clusters ranges between 2 and 4. Because a key part of the analysis is the analysis of the stability across the four years, the same number of clusters is sought for each one of the four years. The optimal number of clusters is selected by using the two indicators discussed above to measure the validity and robustness of the clusters: the Dunn Index and the SD Validity Index.

In 93% of the cases these two indices point to an optimal number of clusters equal to two. Regarding the choice of the optimal algorithm of the three we explored, the findings are conflicting. While the Dunn Index points to the use of the hierarchical method, the SD Validity Index indicates that the medoids algorithm results in more robust results. Therefore, a third indicator was used, the Average Proportion of Non-over-lap (APN). The APN compares two groupings, one based on the full data and a second one based on the full data minus one observation. The APN estimates the average proportion of observations (in our case the PCTs) not grouped in the same cluster. The APN suggests in all cases the most robust algorithm is the hierarchical method.

Consequently, four cluster analyses are estimated, one for each of the years. All clusters are based on the hierarchical algorithm and divide the PCTs into two groups according to the percentage value of the total per-capita expenditure allocated across 20 PBCs (i.e.

the 23 PBCs minus the three we identified above for exclusion). Those PCTs that are consistently in the same cluster during the four years will be grouped. The difference in the expenditure per year and along the four year period of interest will be analysed.

The analysis is done by using the programme R 3.2.0.

A.1.4. Results

A first indication of differences between the PCTs is observed in Table A.1. The range across which the percentage expenditures allocated to each PBC are distributed shows high variability among PCTs. For instance, the expenditures in the PBC `Mental Health Disorders´ is around three time higher in the PCT with the maximum expenditures than the PCT with the minimum expenditure with a standard deviation of over 2.2%. The relative importance of `Mental Health Disorders´ in PCTs' expenditures stands out, with at least 4% more expenditure than that of the second most important PBC `Problems of Circulation.´ Similarly, `Cancers & Tumours´ and `Problems of Circulation´ are also important PBCs in terms of expenditures, showing high variability among PCTs.

The standard deviation from 2008/09 to 2011/12 has kept relatively constant, with a small reduction (around 0.3%) in eight out of the 20 PBCs included in Table .

Table A.1. Percentage expenditure by PBC

	PBC	2008/09				2009/10				2010/11				2011/12			
		Aver.	Max.	Min.	StdDv.	Aver.	Max.	Min.	StdDv.	Aver.	Max.	Min.	StdDv.	Aver.	Max.	Min.	StdDv.
01	Infectious Diseases	1.6	9.8	0.6	1.4	1.7	7.6	0.6	1.2	1.5	5.1	0.6	0.8	1.6	5.6	0.7	0.9
02	Cancers & Tumours	6.2	8.5	3.9	1.0	6.3	8.7	4.3	0.9	5.9	9.6	3.0	1.2	5.9	9.5	3.3	1.0
03	Disorders of Blood	1.3	2.9	0.5	0.4	1.4	2.5	0.5	0.4	1.2	2.4	0.6	0.4	1.2	2.6	0.6	0.4
04	Endocrine, Nutritional and Metabolic Problems	2.8	4.1	1.9	0.4	2.9	3.9	1.8	0.4	3.1	4.3	2.2	0.5	3.2	4.3	2.3	0.4
05	Mental Health Disorders	12.7	22.8	8.8	2.6	12.3	20.5	8.0	2.2	12.2	20.4	6.0	2.6	12.2	19.1	7.1	2.3
06	Problems of Learning Disability	3.6	8.8	0.4	1.2	3.4	9.2	1.2	1.1	3.1	8.2	0.1	1.2	1.7	4.7	0.0	0.9
07	Neurological	4.4	7.5	2.7	0.8	4.4	6.8	2.5	0.7	4.4	6.6	2.4	0.8	4.5	6.6	2.5	0.8
08	Problems of Vision	2.1	3.2	1.0	0.4	2.3	3.7	1.1	0.4	2.4	4.1	1.4	0.4	2.4	3.9	1.3	0.4
09	Problems of Hearing	0.5	1.4	0.1	0.2	0.6	1.4	0.2	0.2	0.5	1.3	0.1	0.2	0.5	1.4	0.1	0.2
10	Problems of Circulation	8.3	11.0	4.8	1.2	8.1	11.4	5.1	1.2	7.5	12.8	4.4	1.2	7.4	9.7	4.4	1.0
11	Problems of the Respiratory System	5.1	7.2	3.4	0.8	5.0	6.7	3.2	0.7	4.7	6.2	2.6	0.7	4.8	6.4	3.1	0.6
12	Dental Problems	4.1	6.6	2.0	0.7	4.0	5.8	2.7	0.6	3.8	5.3	2.7	0.5	3.7	4.9	2.4	0.5
13	Problems of the Gastro Intestinal System	5.1	7.5	2.9	0.7	5.1	7.8	3.2	0.7	4.8	6.3	1.8	0.8	4.9	6.4	3.0	0.7
14	Problems of the Skin	2.1	4.0	1.0	0.4	2.2	4.0	1.2	0.4	2.2	5.0	1.1	0.5	2.2	4.4	1.4	0.5
15	Problems of the Musculoskeletal System	5.1	7.7	2.7	1.1	5.2	8.4	2.8	1.1	5.4	8.3	2.3	1.3	5.6	8.0	2.9	1.1
16	Problems due to Trauma and Injuries	4.0	6.8	0.8	1.0	4.1	7.3	1.0	1.0	3.9	7.1	1.5	0.8	4.0	7.1	2.0	0.6
17	Problems of the Genito Urinary System	4.8	7.2	2.9	0.7	5.0	7.5	3.0	0.7	4.9	7.3	3.4	0.6	5.0	6.9	3.7	0.6
18	Maternity and Reproductive Health	4.0	6.8	1.5	0.9	4.3	8.9	1.5	1.1	3.9	7.3	1.9	1.1	3.9	9.1	1.9	1.1
19	Conditions of Neonates	1.2	3.3	0.4	0.4	1.2	3.3	0.2	0.6	0.9	3.2	0.2	0.4	1.0	2.8	0.1	0.5
20	Adverse Effects and Poisoning	1.2	2.2	0.7	0.2	1.2	1.8	0.6	0.2	1.0	2.3	0.2	0.2	1.0	1.7	0.6	0.2

Source: Authors' elaboration based on data from the Department of Health (2008/09-2009/10-2010/11-2011/12)

In each year, two clusters were estimated. Those PCTs that share the same cluster every year were classified into eight different groups (Table A.2). Each of the eight groups contains at least four PCTs. Seven PCTs were not assigned to any of the eight groups as they did not share the same cluster in every year with any of the other PCTs.

Considerably different patterns of expenditures are presented between Group 6 – 7 - 8 and Groups 1 – 2. Most PCTs are located in Groups 1 and 2. Taken together, these groups account for nearly two thirds of PCTs. These two groups are in the same cluster in the last four periods but in a different cluster in period 2008/09. Similarly, Groups 3

and 4 are located in different clusters in the period 2008/09. It can be expected that these pairs of groups present not equal, but comparable, patterns of expenditures. Other groups that are located in similar clusters are Groups 1 and 3, Groups 2 and 4, Groups 6 and 8, Groups 5 and 8, Groups 4 and 5, Groups 6 and 7.

Table A.2. Number of PCTs per group

	# PCTs	Cluster 2008/09	Cluster 2009/10	Cluster 2010/11	Cluster 2011/12
Group 1	69	2	1	1	1
Group 2	29	1	1	1	1
Group 3	7	2	1	2	1
Group 4	8	1	1	2	1
Group 5	4	1	1	2	2
Group 6	6	1	2	1	2
Group 7	5	1	2	1	1
Group 8	12	1	2	2	2
Not Grouped	7				
Excluded	4				
Total	**151**				

Source: Authors' elaboration based data from the Department of Health (2008/09-2009/10-2010/11-2011/12)

Figure A.2 shows the grouping of the PCTs. While PCTs in Groups 1 and 2 are spread all around the country, the PCTs in Groups 7 and 8 are concentrated mainly in London or in the surrounding areas. The PCTs in Group 8 are all in London. Thus at the very least there may be a 'London effect' that needs to be taken into account.

Figure A.3 and A.4 show, respectively, for 2008/09 and 2011/12, the percentage allocation of per-capita expenditure among the 20 PBCs included in the analysis. Groups 5 and 8 invest a considerably higher proportion of their expenditures in `Mental Health disorders´ in comparison with the other groups. In the case of Group 6, the investment in `Mental Health disorders´ is similar to Groups 5 and 8 in the period 2008/09, however, a decrease can be observed in 2011/12, which relocates this Group to a level of expenditure similar to Groups 2 and 3.

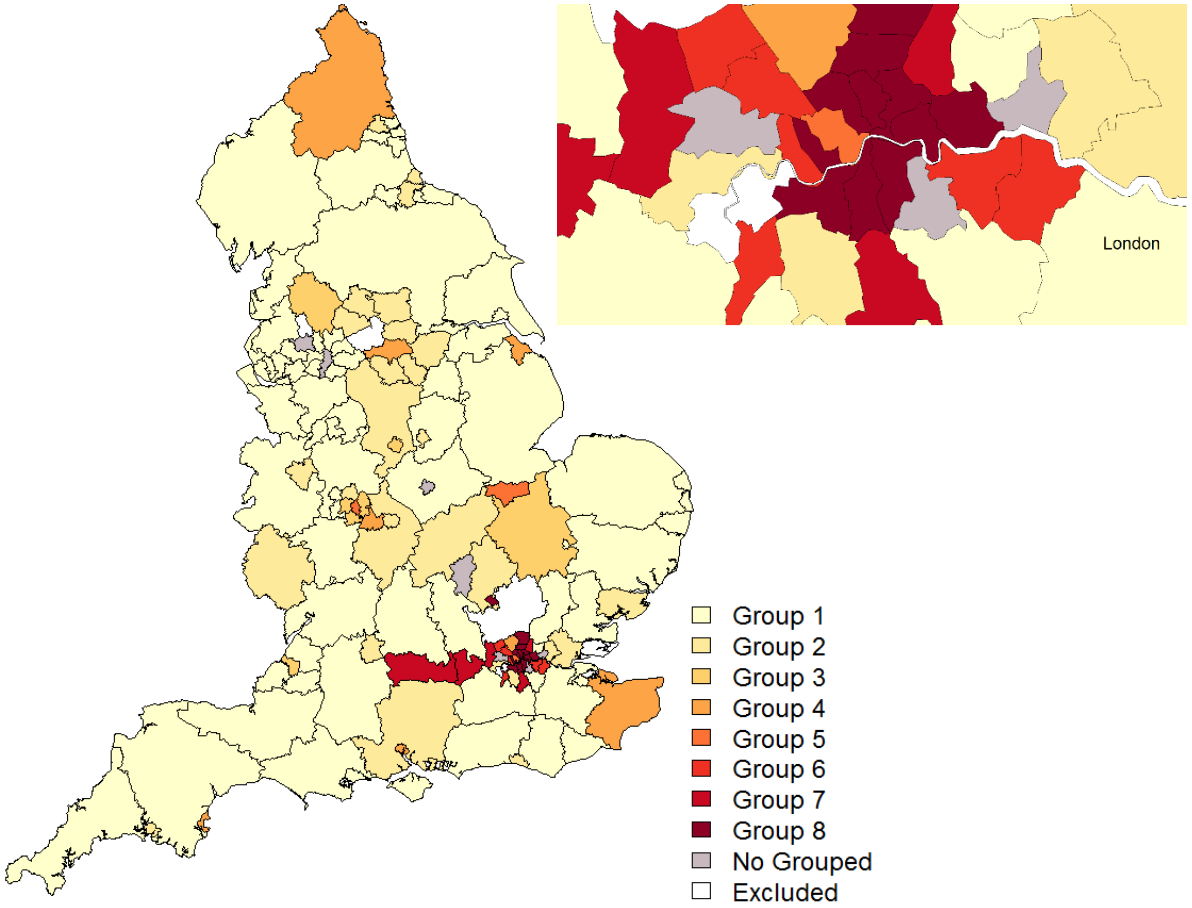
For the PBC `Maternity and Reproductive Health, it can be observed that the last three groups (Groups 6, 7 and 8) spend a higher per-capita proportion than the other groups, particularly in the period 2011/12.

Group 1 is characterized particularly by investing a higher proportion in circulatory and of gastro intestinal diseases in comparison with the other groups. Group 8 spends the highest proportion in the PBCs `Infectious Diseases´ and `Conditions of Neonates´.

Groups 5 and 8 spend a relatively smaller proportion of the per-capita health expenditure on `Cancers & Tumours´, `Problems of the Musculoskeletal System´ and `Neurological´.

As mentioned above, the expenditures patterns of Groups 1 and 2 differ only in the first year of analysis, 2008/09. For instance, Group 1 dedicated a higher proportion to tackling problems related to the circulatory system and to learning disability than Group 2.

Figure A.2. Cluster Analysis



Source: Authors' elaboration based data from the Department of Health (2008/09-2009/10-2010/11-2011/12)

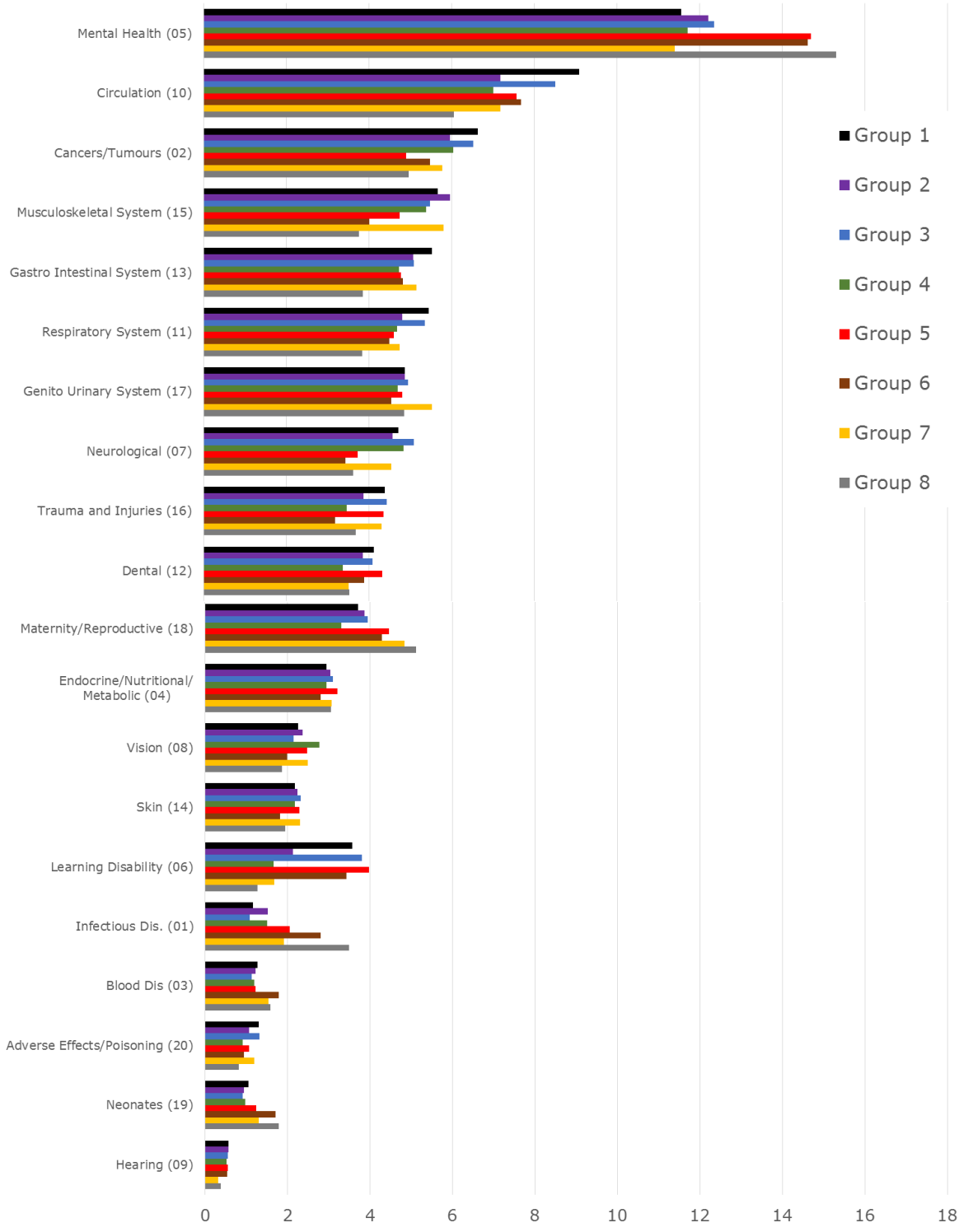
For the PBC `Problems of Learning Disability´ there is a considerable reduction over the period of analysis in the share of spend invested by Groups 1, 3, 5 and 6 , from a value close to 4% in 2008/09 to a value of under 2% in 2011/12.

A.1.5. Conclusions

This analysis indicates that English local health care commissioners (PCTs) cannot all be considered to behave similarly to one another. This because there is more than one cluster in each year and there exist a number of PCTs that remain in the same distinct cluster over time.

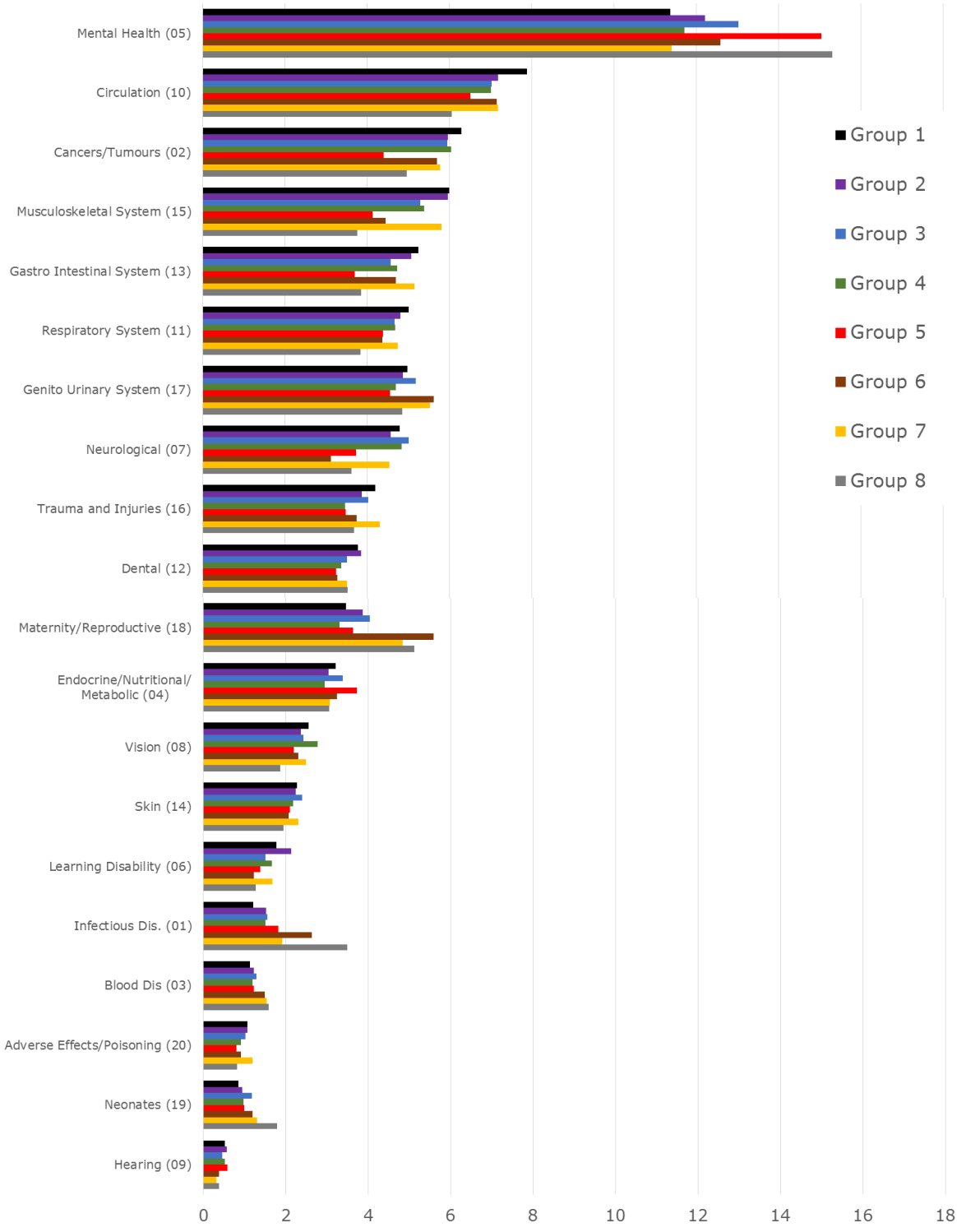
The analysis indicates that in each year at least two significant clusters exists. After identifying those PCTs that share the same cluster, we were able to split the PCTs into eight different Groups. Some of the Groups are similar to each other, since their members have shared the same cluster in some periods. We observed two clearly different patterns of behaviour - one followed by the PCTs in Groups 1 and 2 and one followed by Groups 5, 6, and 8. The PCTs in Groups 5, 6 and 8 are mainly in the London area. Nevertheless, even when some Groups share similar clusters, they still show differences in the evolution of their expenditures over the four year period of analysis, for instance, in the percentage dedicated to ´Mental health disorders´ and `Problems of learning disability´. This suggests the possibility that the PCTs in different Groups could react differently in response to a need for a reduction in expenditures.

Figure A.3. Percentage per-capita expenditure 2008/09. Average per group.



Source: Authors' elaboration based data from the Department of Health (2008/09-2009/10-2010/11-2011/12)

Figure A.4. Percentage per-capita expenditure 2011/12. Average per group



Source: Authors' elaboration based data from the Department of Health (2008/09-2009/10-2010/11-2011/12)