Healthcare funding and its relationship to equity and outcomes: A Qualitative Comparative Analysis of Commonwealth Fund and OECD data

Abstract

This paper examines OECD data to explore the relationship between the level and means of funding of 11 different healthcare systems on the one hand, and Commonwealth Fund data which assess overall system performance, equity and health outcome measures on the other. It utilises Qualitative Comparative Analysis (QCA), and the idea of 'fitness landscapes' to analyse combinations of funding combinations and outcomes. It finds that health systems with relatively high levels of voluntary health insurance tend to be in sets of countries with low outcome measures, but that high outcome achievement is often part of fitness landscapes that are dependent upon specific combinations of funding patterns that make generalisations more difficult. The paper finds that medium levels of out-of-pocket expenditures, in combinations with other specific funding factors, and in opposition to existing work, are not incompatible with both high equity and high health outcomes scores. Finally, it explores how 'contradictory cases' can be used to explore how systems falling short of the outcomes of others with the same funding patterns might find improvements.

Introduction

The funding of healthcare systems remains a live political issue. Most developed nations spending around 10% of GDP on healthcare, and countries such as the US considerably more than that. Healthcare expenditures also appear to have an inexorable tendency to increase as a proportion of GDP over time (Johnson et al., 2017). Even as governments put increased pressure on health systems to achieve greater efficiencies, demand for health services continues to rise.

Research on health funding is often based on typologies involving the degree of

1

public control over them (Blank et al., 2018: chapter 3), but here we present an analysis that combines both the sources and levels of funding, as well as relating it to service outcomes. While we know that funding is a key area of health governance, and that countries have different strategies for how they pay for healthcare, we know less about how levels of total healthcare expenditure combine with the sources of healthcare funding. Perhaps most importantly of all, we do not know what the relationship is between different funding combinations and health system outcomes.

Existing typologies of health systems are often organised around public and private expenditures, or around public, social insurance and private expenditures (Wendt et al., 2009). However, examining funding according to the OECD's main statistical categories of public expenditure, out-of-pocket expenditure, and voluntary health insurance, as well as through their funding levels, allows us to get past definitional difficulties between public funding and social insurance, and between different types of private funding methods. Using the OECD categories also allows to revisit existing typologies by evaluating their outcome results as well as their health governance ideal type, to question, for example, as to whether the UK, New Zealand and Sweden fit within the same type of publicly-financed health governance.

Exploring health systems through their funding methods clearly omits important factors such as the staff employed, and the level of technology in different health systems. Those factors can form the basis of future work. It is crucial, however, that we understand not only how different funding systems combine in health systems, but also what outcomes those funding systems lead to. To that end, the paper utilises the Commonwealth Fund's 2017 ratings of health systems, which are broadly representative of those in developed economies, and which include countries from across the world with a range of funding patterns and outcomes. As such, although using the Commonwealth Fund data limits our analysis to eleven countries, it provides us with a robust and consistent dataset of health system

outcomes to explore in relation to different funding patterns.

To explore the relationship between funding and outcomes, the paper uses Qualitative Comparative Analysis (QCA), which is ideally suited to exploring relationships between non-independent causal factors and outcomes – being specifically designed by its founder Ragin to that end. QCA is a highly robust method, but still little-used in the analysis of health systems, giving the paper's method, as well as its findings, a clear originality.

The paper finds that funding health systems with high levels of voluntary health insurance appear to perform relatively poorly across the dimensions explored here, but also that moderate out-of-pocket expenditures are not incompatible with strong health outcomes or equity scores.. The paper also makes a methodological contribution in its use of QCA, not only in its application in a novel area, but in its exploration of 'contradictory' cases, which are defined and explored in depth below.

Health financing, funding and outcomes

Despite research suggesting a relationship between the level of health funding and health outcomes (Deaton, 2013), in developed countries at least, this relationship does not always hold (Cremieux et al, 1999). Countries such as Australia with comparatively low overall health expenditures have amongst the best health outcomes in the world, while the country with the highest expenditures of all, the United States, has some of the worst. Considering how different funding mechanisms combine to produce different health system outcomes is clearly an important topic worth exploring in greater depth.

There are several possible measures of total health expenditure. Here we will use total health expenditure as a percentage of GDP because of the difficulties of comparing expenditure per capita between different countries, even if we take into account attempts to measure purchasing power parity.

Chart one shows total health expenditure as a proportion of GDP for the 11 countries included in the Commonwealth Fund measures of outcomes.

Figure one – total health expenditure as a % of GDP

Total Health Expenditure as a % of GDP
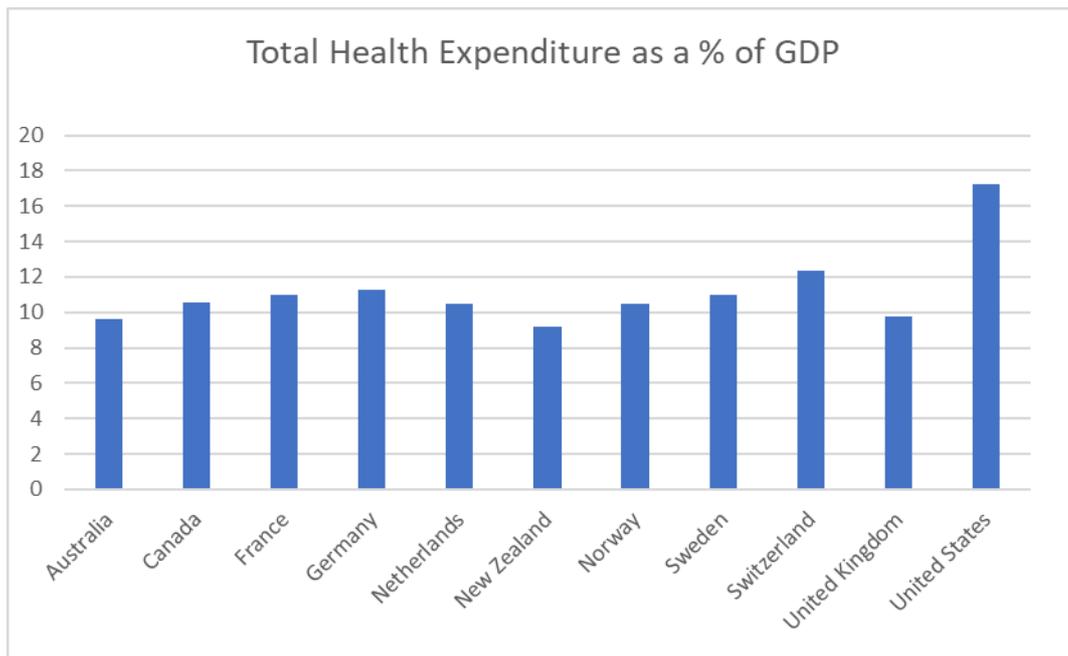
From this we can see the US is by far the largest total spender on healthcare, with New Zealand and the UK being the lowest. There are also significant differences between expenditure levels, especially if we take into account that these gaps have cumulative effects. In order to explore health funding, we must first be clear about what we are measuring – here we measure total health expenditure, but also how that expenditure is funded. Different methods of health financing are often the basis of typologies of health systems, but outside of studies of comparative health systems as a whole (Johnson et al., 2017), or research generating typologies of health systems (Marmot and Wendt, 2012), there is little work exploring what effects we might expect between health financing and system outcomes (although see Schieber et al., 1991; Wendt, 2015). Existing work does, however, offer some clues as to the dynamics we might expect to find.

OECD health statistics[1] explore the levels of funding but also its sources, with its top-

---

[1] Available at http://www.oecd.org/els/health-systems/health-data.htm

level measures locating all healthcare systems in hybrid forms which combine three funding sources . First, government funding provides a healthcare safety net (Abel-Smith, 1994), as well as providing crucial clues as to the overall size of the health system and the role of government within it (Blank et al., 2018: chapter 3). Because of the cost of healthcare, government funding is a significant component of all the systems explored here to varying degrees.

Health systems with proportionally high levels of public expenditure should achieve greater equity and efficiency. Even sceptics of public services such as Lawson (1991) came to conclude that funding healthcare through the public purse created an efficient system through greater control of expenditure, and higher levels of government funding should also improve system equity if they are used to alleviate affordability problems. However, Berger and Messer (2002), find that 'higher mortality rates are observed in countries and years in which more of the expenditures are publicly financed' (p. 2112), suggesting that the link between government financing and outcomes may not be as straightforward as we might first suspect.

The second OECD category of health funding is through 'out-of-pocket' expenses, in which people pay for (or at least contribute to) the costs of healthcare from their own private funds. This is often presented as a means of attempting to increase personal responsibility (Carrera et al., 2008). Out-of-pocket expenditures, where they are used, create the ability for public funding to be supplemented by those who can afford to do so, potentially increasing the total spend on healthcare. Increasing total expenditure in this way may act to improve health outcomes for those that can afford additional services, but there may be equity implications for those that cannot. We would expect to see in countries such as Switzerland, which has high out-of-pocket expenditure as well as relatively high overall expenditure, to achieve strong health outcomes (at least for those who can afford to access healthcare), but

perhaps with an equity trade-off for those that cannot afford the extra payments. As Marmot (2007) perhaps makes clearest 'Out of pocket expenditures for health care deter poorer people from using services, leading to untreated morbidity' as the 'larger the proportion of health care that is paid out of pocket, the larger the proportion of households that are faced with catastrophic health expenditures' (p. 1159).

The third OECD funding mechanism, that of voluntary health insurance, utilises insurance policies to cover the costs of healthcare, and which are paid for by individuals or by organizations on their behalf. In common with out-of-pocket expenditures this may emphasise personal responsibility (Blank et al., 2018: chapter 3). The most obvious example of a country making significant use of voluntary health insurance is the United States. However, the use of voluntary health insurance has well-known issues around the principal-agent problem where insurers have to meet the costs of care, but it is individual members of the public who drive the demand for health service (Smith et al., 1997). The use of voluntary health insurance creates complex incentive relationships, especially where insurance is used as a means of supplementing pay. In such systems, total health costs may be hidden from those receiving it as pay, making overall expenditure difficult to control (Blank et al., 2018: chapter 3), and creating equity issues if some groups are unable to afford health insurance and are not covered by public funding.

Patterns between the means and level of funding and measured outcomes often leads to trade-offs between these different dimensions, especially in terms of how they impact on measured outcomes in terms of health equity and health outcomes (Blank et al., 2018: chapter 8). Dominant ideas around personal responsibility, which is a dominant theme in US health policy discourse, puts an emphasis on voluntary insurance and out-of-pocket payments (Starr, 2013), while advocates of taking a social determinants approach, instead emphasise wider societal factors in health systems, and the need for more publicly co-ordinated responses

(Marmot, 2015). When we explore these trade-offs in relation to specific outcomes, the notion of a 'fitness landscape' is a useful one, so it is to this topic that we now turn.

Fitness landscapes

The concept of a fitness landscape originates in evolutionary biology where it is used to visualise the relationship between genotypes and their reproductive success, with higher fitness being shown through the achievement of a higher point on the landscape (Dennett, 1996). Genotypes are mapped in fitness landscapes according to their similarities or differences in achieving outcomes, and similarly, we might think of the four funding factors as providing different combinations which position them on the landscape, and with those positions fitting with different levels of measured outcomes. Moving from one fitness 'peak' to another will involve the need to adapt to using different mixes of factors (Mischen and Jackson, 2008), but also raise the possibility of significant disruption occurring as a result of such change, with the system having to move away from one fitness peak into a 'valley' between before improvement can occur again.

Fitness landscapes provide us with a means not only of identifying which patterns of funding combine in relation to different outcomes, but also alert us to look for cases which appear to have promising funding patterns, but which do not achieve successful outcomes – or 'contradictory' cases. We can use these insights to explore similarities and differences in funding patterns to try and explain these patterns, and to try and assess how improvements might be made. Fitness landscapes are also helpful in identifying that, once a particular combination of funding factors is in place in a healthcare system, it may be extremely difficult to change it. Government funding in an area such as healthcare is path-dependent (Jessop, 2002), in that it creates a range of positive feedback mechanisms that are difficult to adjust once in place (Pierson, 2004). Institutional arrangements quickly become self-sustaining with interest groups in place which are prepared to defend them (Kingdon, 2011:

appendix), even should the system itself appear to be performing badly (Gordon, 2009). Moving the funding of a health system onto a more successful basis might require a wholesale reform of current funding arrangements with all the disruption and risk that can bring, but, if the current system is only a small movement along a single causal factor from more successful systems, could be more straightforward.

By exploring health funding through the lens of combinatory causation and fitness landscapes allows us to begin to assess the outcomes of different health systems in relation to their funding, and the scale of the changes that they might need to make to reach a higher fitness point, or to move to another fitness point with a different configuration of factors.

Method

Statistical measures are the tools used to explore how different causal factors work together to produce different outcomes. However, with causal factors that are not independent of one another, it becomes difficult to control for the effects of each factor, and it is beneficial to look at how the different factors relate to one another in ways in combinatory forms instead.

Qualitative Comparative Analysis (QCA) attempts to meet these challenges. QCA is an approach first developed by Ragin (2008; 2014), and which continues to be taken forward by him, along with a range of other innovators of the method (Dusa, 2018). It is arguably most developed in the field of comparative politics (Vis, 2009), with its strength being in fields where causality is complex (Hudson and Kuhner, 2010). Here we use the fuzzy-set version of QCA because of the extra precision it brings to analysis compared to the crisp-set variant, with both causal factors and outcomes being calibrated on a scale between 0 and 1, rather than falling on these two crisp set analysis poles. A full description of the means of calibration for the fuzzy sets is presented below.

QCA is suited to the project of examining how different funding systems affect different outcomes because of its concern with multiple causation pathways and equifinality. Here we have a dataset where some aspects (type of funding) are logically dependent on one another, and in which health systems make use of very different combinations of these different funding types.

Data outcome measures

Given the extent of debate concerning healthcare funding, it is perhaps not surprising that there are a range of different systems for measuring the outcomes of health systems. Different methods for measuring how well healthcare systems are doing might assess their outcomes through composite indexes, or create measures that attempt to examine a range of different health and healthcare dimensions.

The Commonwealth Fund measures of healthcare performance have been in place since 2004, being published every few years, and works across a range of different dimensions. The most recent report, published in 2017 (Schneider et al., 2017) explores 11 developed countries in terms of their care process, access, administrative efficiency, equity and outcomes, and covers a range of countries with different sources of funding and outcome levels. No system of measuring healthcare system outcomes is perfect. The Commonwealth Fund system has strengths in assessing a diverse range of systems across a range of different dimensions in a way which is transparent by giving the data they are based upon in the report. It is therefore a credible source for the kind of analysis being conducted here.

In this paper we will use the 'overall' measure in the data, despite the limitation of it being a simple mean of the other measures, to consider the question of how different funding mechanisms link to overall health system outcomes. However, the main focus will be on the 'outcome' and 'equity' measures because of their importance in the discussion above about the trade-offs involved in different levels and means of healthcare financing. Finally, the

paper explores the relationship between outcomes *and* equity – asking whether healthcare systems can achieve strong performances on both measures at the same time, when the two can often be in tension to one another.

Data calibration

In QCA it is crucial that data is calibrated against clear criteria to position it within the fuzzy set in an appropriate way. The ideal is that such calibration occurs against an external standard. While using the outcome measures from the Commonwealth Fund, the results have been standardised in relation to the other countries in the sample. As there is no accepted external standard for measures of health outcomes and equity, it seems unavoidable but to derive our calibrations using it. The report scores were converted into fuzzy sets by each outcome measure being plotted (in line with (Ragin, 2008)), and clusters of scores being derived visually to ensure that they were located within the appropriate range of similar fuzzy scores –especially playing close attention to selection of the crucial cross-over point of 0.5. Finally, once the cross-over point, maximum and minimum had been selected, the data was scaled using the direct method (Ragin, 2008).

Causal factors (funding sources and level) were calibrated using a different process. Here data on funding measures from the OECD was used, and each factor was initially graphed to look for clusters to identify which countries should fall 'in', or 'out' of the set, and a cross-over point (0.5) between the two selected based on substantive knowledge of the cases. For GDP funding, for example, the US clearly represents the highest spend, and New Zealand the lowest, but with the UK NHS also generally presented as having lower spending (with the rise in the 2000s being offset by falls since (Greener, 2018)), and France and Germany being characterised as having higher spending levels, especially when those cumulative differences are taken into account, so are above the 0.5 threshold. Different calibration methods were then applied, including the indirect method (Ragin, 2008) on a four-

point and six-point scale, before the sets were finally calibrated using the direct method, and the results of the different calibrations compared. The final results below are those based on the direct process, with only minor differences coming from the three different calibrations. The process yielded the following calibrated data set:

Table one – Calibrated data set

Country	GDP	Gov	Pocket	Vol	Overall	Equity	Outcome	EquOut (Equity AND Outcome)
Australia	0.12	0.26	0.76	0.51	0.95	0.38	0.95	0.38
Canada	0.52	0.32	0.52	0.56	0.25	0.2	0.19	0.19
France	0.56	0.63	0.05	0.54	0.13	0.14	0.77	0.14
Germany	0.59	0.94	0.34	0.09	0.65	0.51	0.32	0.32
Netherlands	0.51	0.79	0.31	0.2	0.91	0.84	0.54	0.54
New Zealand	0.05	0.74	0.35	0.22	0.75	0.3	0.37	0.3
Norway	0.51	0.95	0.51	0.05	0.75	0.63	0.89	0.63
Sweden	0.57	0.92	0.55	0.06	0.66	0.79	0.94	0.79
Switzerland	0.71	0.19	0.95	0.22	0.66	0.78	0.84	0.78
United Kingdom	0.16	0.66	0.53	0.15	0.96	0.95	0.07	0.07
United States	0.95	0.05	0.22	0.95	0.05	0.05	0.05	0.05

Data Analysis

The data was analysed using 'enhanced standard analysis' as outlined by Schneider and Wagemann (2012) and Dusa (2018). We carefully reviewed counterfactuals for inclusion on the basis of their credibility, moving back and forward between the data and its emerging findings in an iterative way, and as recommended by Ragin (2008) to avoid the mechanical interpretation of the data set. The findings presented here are robust in that they appear across a range of reasonable assumptions about calibration (see above) and inclusion of relevant counterfactuals to be reliable conclusions from QCA analysis, while balancing the need for both consistency and coverage in the data.

The data analysis was carried out in Ragin's fsQCA 3.0 software, as well as in R using the 'QCA' package (Dusa, 2018) in order to cross-check results. fsQCA allows us to more easily explore the consequences of using different tied prime implicants, whereas the QCA R package allows us to more easily to identify remainders (counterfactuals) which should and should not be included than by using fsQCA.

Finally, we looked especially for cases which appeared to fit with what appeared to have patterns of funding which achieved the consistency thresholds for strong outcomes, but which did not achieve those outcomes in their own results. These are 'contradictory' cases, and require greater thought and explanation involving moving back and forth between the data and cases as Ragin (2008) suggests.

Results

Overall health system ratings

We first looked for patterns of funding that were linked to a high overall score for the performance of healthcare systems. No single causal factor was a necessary condition for a high overall system rating. For sufficient solutions, we included directional expectations that higher health expenditure would lead to higher overall ratings, based on existing research from Deaton (2015) although acknowledging the weakness of the United States in terms of this assumption. Not including this directional expectation leads to marginally different results, but the same substantial underlying analysis. The following solution resulted:

Solution term	Raw Cov.	Unique Cov.	Cons.	Cases
Gov*~Vol	0.76	0.33	0.82	NOR,SWE,GER,NETH,NZ,UK
GDP*Pocket*~Vol	0.45	0.07	0.93	SWI, SWE, NOR
~GDP*~Gov*Pocket	0.27	0.04	0.89	AUS

The overall solution has a coverage of 0.87 and a consistency of 0.81.

The first solution term is a combination of high Government funding (Gov) and low Voluntary Health Insurance (~Vol). This is the solution with the highest unique coverage (0.33) of the overall solution. The second term is a combination of high expenditure (GDP)

with high out-of-pocket expenditure (Pocket), but low voluntary insurance expenditures. This solution provides very little unique coverage in terms of the solution (0.07).

The third solution term is perhaps a little unexpected given the first two – comprising low GDP expenditure, low government expenditure (both factors of which are in the high sets in the first two solution terms), and high out-of-pocket expenditure. This combination represents Australia only in this solution – but is still an important addition because that country's health system comes second in the overall rankings.

The three solutions here give us three initial 'fitness peaks' emerging for high overall performance, the first of which is based around high government expenditure combined with low voluntary health insurance. If we were to pick one funding solution term that most 'covers' overall high performance, it would be this one because of its higher unique coverage score compared to the other solutions. This combination covers six of the eight countries with high overall scores. We can think of this as being the high government expenditure fitness peak, which seems to combine strongly with countries having low voluntary health insurance.

The second fitness peak is based on high out-of-pocket expenditure, which replaces high government expenditure, and as well as sharing low voluntary health insurance with the first, adds high current expenditure. This second solution term brings in Switzerland into its coverage, and has Norway and Sweden in common with the first term. However, its low unique coverage of the solution should give us initial pause in regarding it as adding much to our first solution.

The third solution term covers Australia only, but which is an important addition as Australia are ranked second in the overall healthcare ratings. This can be considered either as a variant on the second fitness peak (as it also shared high out-of-pocket expenditure), but equally as one based on low GDP expenditure (contrary to our initial expectations), or low government spending, both of which are, as mentioned above, causal factors running in the

opposite directions as the first two solution terms. That these factors run in opposite directions emphasises Australia's uniqueness in achieving strong overall system rankings.

Finally, if we explore cases of low overall rating achievement, then two solutions terms appear with a combined consistency of 0.90 and a coverage of 0.62. First, there is a combination of high expenditure, low government expenditure and high voluntary health insurance, with a consistency rating of 0.89 and includes the cases of the US and Canada. Second, a combination of high expenditure, low out-of-pocket and high voluntary health insurance also has a consistency of 0.89 and includes the case of the US (again) and France. These low overall score solutions suggest that when high expenditure and high voluntary health insurance levels (common to both solution terms) combine with either low out-of-pocket expenditures or low out-of-pocket expenses, then this is not a good combination for the overall performance of health systems.

Outcomes

The Commonwealth outcome measure combines scores from data three categories: population health outcomes (including chronic disease and mortality of populations); mortality amenable to health care (including deaths under age 75 from specific causes that are considered preventable); and disease-specific health outcomes measures (including mortality rates following stroke or heart attack and the duration of survival after a cancer diagnosis).

Moving from the overall ratings to outcome rankings, the biggest country change concerns the UK – which moves from first overall to second-worst. Australia moves from second overall to the highest outcome score. Sweden, Norway and Switzerland continue to do well, with the Netherlands, New Zealand and Germany falling back a little. Canada and the US perform poorly on both measures. For analysis, we again included directional

expectations in terms of higher GDP leading to higher health outcomes (even bearing in mind the case of the USA) on the grounds that higher expenditures, at least in principle, allowed a greater scope for services to be provided.

There were no single causal factors that were necessary for high outcome achievement. In terms of sufficient solutions, a more complex range of terms emerge compared to the overall system measures:

Solution terms	Raw cov.	Unique cov.	Cons.	Cases
Gov*~Vol	0.72	0.16	0.68	NOR, SWE, GER, NETH, NZ, UK
GDP*Gov*~Pocket	0.48	0.02	0.85	GER, FRA, NETH
GDP*Pocket*~Vol	0.49	0.08	0.89	SWI, SWE, NOR
~GDP*~Gov*Pocket*Vol	0.28	0.04	0.82	AUS

The overall solution has a consistency of 0.70 and a coverage of 0.87.

The first solution term (which has the highest unique coverage) combines high government expenditure with low voluntary insurance. It is the direct equivalent of the first overall health system solution term, and covers the same six countries. However, in the overall ratings, all the countries were in the set of highly-performing countries, whereas here the UK, Germany and New Zealand all have a score which puts them in the low set of outcomes performers – they are therefore contradictory cases – a topic we will pick up in the conclusion. This explains the consistency of this solution term being 0.68, even though the cut-off for the solution as a whole was 0.78.

The second solution term combines high GDP expenditure with high government spending and low out-of-pocket expenditure. However, it covers only one unequivocally

strong outcomes nation – France, with New Zealand only just being part of the set of high outcome countries, and Germany in the set of low outcome scorers – and so being a contradictory case in this solution set as well as in the first solution term.

The third soluion term is one of high government expenditure with high out-of-pocket expenditure and low voluntary health insurance. It covers the second, third and fourth-ranked outcomes nations, and so is clearly substantively important even though the unique coverage of the solution is only 0.08.

Finally, we have an additional combination of low GDP with low government expenditure, with high out-of-pocket and high voluntary health insurance. This solution again represents Australia, based on a pattern of funding similar to the third solution term for overall health systems, but adding high voluntary health insurance to it. This addition again emphasises the rather unique characteristics of the Australian system.

Our fitness landscapes for health outcomes present, again, the distinctive case of Australia (solution term four) based on low GDP and low government expenditure, but high out-of-pocket and high health-insurance expenditure. The second landscape combines high GDP with high out-of-pocket expenditure and low voluntary health insurance (solution term 3) - and covers three of highest outcomes scorers - Switzerland, Norway and Sweden, so we can regard it as being probably the most substantively important.

The third landscape is based on high government expenditure, but is more complex. One combination (solution term 2) comprises high health expenditure with high government expenditure and low out-of-pocket expenditure, and so include France of the top performers, but includes Germany as a contradictory case. The other combination (solution term one) in the landscape combines high government expenditure with low voluntary health insurance, but introduces the contradictory cases of the UK, Germany and NZ.

In terms of low overall ratings, a single solution of high expenditures with low

government expenditure and high levels of voluntary insurance has a consistency of 0.90 and includes the cases of the US and Canada. This solution adds to a sense that high overall ratings and high outcomes seem (except in Australia) to be linked to low levels of voluntary health insurance, whereas high levels of voluntary health insurance appear in solutions for low overall ratings and low outcomes ratings. High levels of voluntary health insurance also appear alongside high expenditure in solution terms for low overall system ratings and low outcome cases, perhaps indicating a loss of control over health expenditure as a result of the extensive use of voluntary health insurance.

Equity

The Commonwealth Fund equity measure incorporates 11 survey measures including service affordability, access and service co-ordination, and which assess, for each of those measures, the percentage-point difference between the two groups based on the reported income of above-average vs. below-average relative to the country's median income. For health equity, low voluntary health insurance is a necessary relation with high consistency (technically inclusion) and relevance:

Figure two – X,Y chart of ~Vol (Negvol) versus Equity
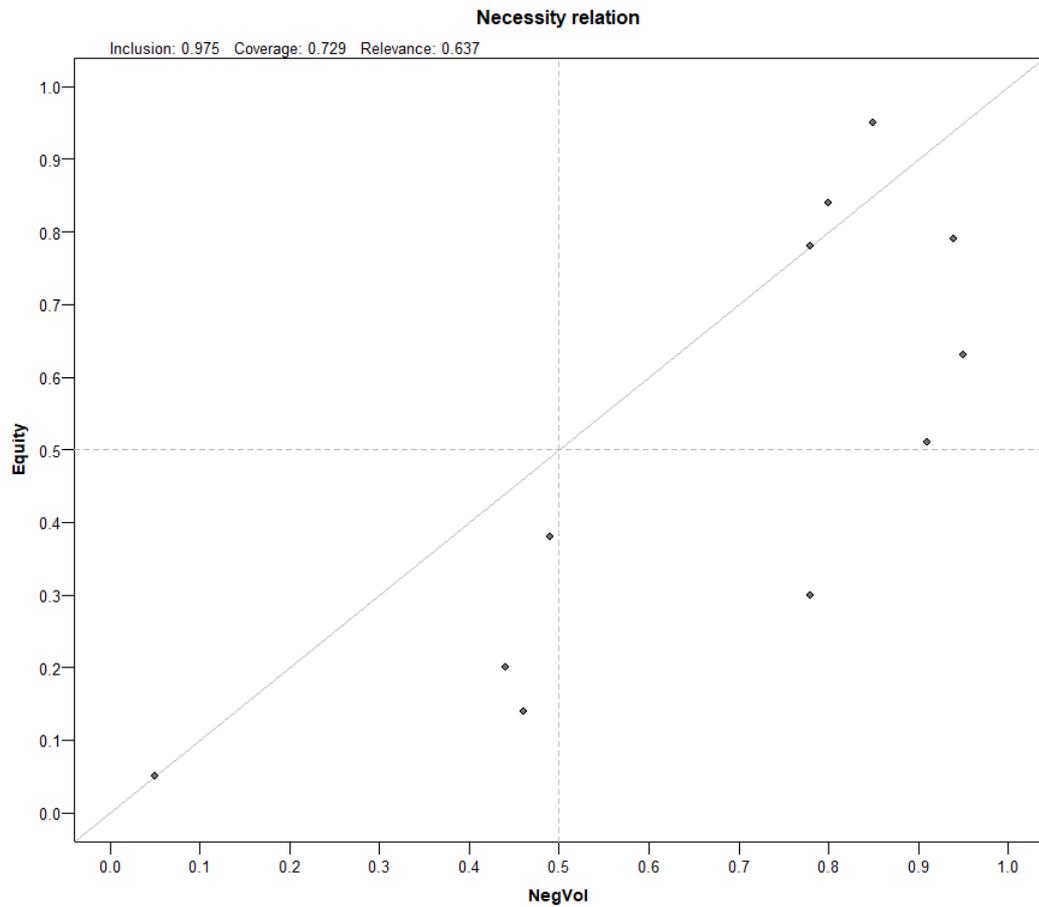
In calculating the sufficient solution for equity, we included negative directional expectations in respect of voluntary health insurance and out-of-pocket expenditures on the grounds that these were the elements of health funding that those with limited resources would be least likely to be able to afford. This resulted in the following solution:

Solution terms	Raw cov.	Unique cov.	Cons.	Cases
Gov*~Vol	0.81	0.27	0.72	NOR, SWE, GER, NETH, NZ, UK
GDP*~Vol	0.63	0.09	0.85	SWI, GER, SWE, NETH, NOR

The overall solution here has a coverage of 0.90 and a consistency of 0.73.

The first solution term combines high government expenditure with low voluntary health insurance. This combination has the highest unique coverage amongst the solution terms. Of the countries in the set with this combination, New Zealand represents an contradictory case, but the other countries cover five of the top six scores for the category. This is therefore both quantitatively (in terms of unique coverage) and substantively the most significant solution term for equity.

The second combination is of high GDP with low voluntary health insurance. It adds Switzerland to the solution for term one, but loses the UK (the highest equity scorer) along with New Zealand (the contradictory case).

Here then, the dominant high equity fitness landscape term is one of low voluntary health insurance, combined with either high government spending or high GDP and high out-of-pocket expenditure.

The solution for low equity scores gives two clusters in its solution – one for high expenditure and high voluntary health insurance (0.95 consistency), and covering the cases of the USA, France and Canada. The second cluster combines low government expenditure with high out-of-pocket expenditure and high voluntary health insurance (0.94 consistency). This includes the cases of Canada (again) and Australia. Both these solutions have high

voluntary health insurance in common, emphasising its link with low equity achievement.

Equity AND outcomes

Finally, we can examine the combination of equity and outcomes to explore which patterns of funding best achieve this. In QCA, the 'AND' operator takes the minimum value of the two numbers, rather than the mean. Here we used directional expectations of low voluntary health insurance and high government spending based on the findings from analysis above, but with the solution below proving resilient to other credible directional expectations as well.

This more challenging criterion leads to only four countries having a score above 0.5 - the cross-over point for a high outcome. Applying enhanced standard analysis gives one combination only:

Solution term.	Raw Cov	Unique Cov	Cons	Cases
GDP*Pocket*~Vol	0.70	0.70	0.89	SWI, SWE, NOR

This solution covers three of the four cases with high equity and outcomes scores, with the Netherlands only not forming part of the solution term. It is made up of high spending, high out-of-pocket expenditure, and low voluntary health insurance. This solution, then, gives us a clear understanding of the pattern of funding sufficient to achieve both high equity and outcomes.

Discussion

Health systems with patterns of funding based on high levels of voluntary health insurance do not appear to be a strong route to achieving strong health system performance,

equity or outcomes. The countries with high levels of voluntary health expenditure insurance tend to do poorly across the different measures here. Canada, France and the United States are consistently poor performers across the measures used here (France achieves good results on outcomes only). It would seem then, in designing a health system, high levels of voluntary health insurance do not lead to strong health system outcomes. The exceptional case here is Australia, which only just falls into this category (0.51), and where its voluntary health insurance level forms part of a very distinctive combination of other funding factors.

Australia's particular pattern of funding of low GDP combined with low government expenditure, but high out-of-pocket expenditure and high voluntary health insurance, seems to be linked to strong overall health system performance, as well as with strong health outcomes, but does less well in terms of equity. We might expect Australia to achieve less in terms of equity based on the exploration of existing research in the first part of the paper, but also suggest that addressing the relatively low government expenditure could be a means of redressing the equity shortfall this causal combination is subject to. It is interesting, however, that Australia is the only country with this combination of funding type and level in the Commonwealth Fund sample, especially as it seems to lead to lower overall health expenditures, but high outcomes –even if there seem to be equity problems generated as a result.

It is surprising that Switzerland scores well in terms of equity, given its very high use of out-of-pocket expenditure. If we drill down into the Commonwealth Fund data in more depth on this point, however, we find that cost is an issue in terms of equity for the Swiss, with 30% of people (second only to the 44% of the US) reporting cost-related access problems to healthcare in the last year, and 14% (second only to the US score of 18%) having serious problems paying medical bills. These aspects of affordability are clearly important, but as Switzerland scores highly on most aspects of access (another key component of the

equity measure) this compensates for them in the overall equity measure. These affordability problems cast some doubt on the Commonwealth fund equity measure as using high out-of-pocket expenditures on the poorest deciles is likely to lead to an increase unmet health needs (Kaminska and Wulfgramm, 2018). Taking this more fine-grained approach to the data points to possible system improvements, even for the high-performing Switzerland, asks us instead whether Norway and Sweden might present a more promising funding combination instead.

If non-public modes of funding such as voluntary health insurance and out-of-pocket expenditure are meant to increase personal responsibility, we might expect to see an improvement in health outcomes for countries making use of those means of health financing. However, as noted above, this is certainly not the case for voluntary health insurance, but out-of-pocket expenditures do seem to form the basis for a fitness landscape under specific combinations with other factors – a finding to which we return below.

Conclusion

Exploring how different levels of funding fit with different means of funding, and how those combinations fit with different levels of health system outcomes, give us key insights in health system performance, despite the limitations of focusing on these causal factors alone. The paper's analysis gives an empirically-grounded set of fitness landscapes which can form the basis for further work which might incorporate them into the analysis of health systems.

The picture presented in the paper is complex. It involves us taking into account both funding levels and the means of funding, and identifying combinations that resemble complex fitness landscapes based on the outcome measure which is our focus. This helps us understand which combinations work better than others, as well as offering a means of identifying cases which can improve on their current levels of performance. There are two

main possibilities for such improvement based on the data here.

First, contradictory cases which are associated with patterns of funding for which other countries are performing better, can lead us to explore case differences to explore what might be causing such inconsistencies and find possible means of outcome improvement. Second, and relatedly, where cases are configured in a way in which one causal factor is different from successful countries, we might suggest that it is worth exploring what difference changing that factor might make – effectively forming a counterfactual analysis.

Switzerland is perhaps the nearest we have to the first possibility, where we noted that, despite it being one of the three nations achieving both high equity and high outcomes scores, alongside Sweden and Norway, a more nuanced exploration of the data underlying the Commonwealth results raised concerns about the affordability of the Swiss health system, and so to an inconsistency with the other two cases. Although Switzerland was not a contradictory case, going back-and-forth between the cases and the data led to greater insight as to its status in relation to Sweden and Norway, and to us questioning the equity measure in the Commonwealth Fund report in this instance. Exploring the case in this way, allows us see the potential for improvement in equity for Switzerland were it to move to a system with higher government funding (provided that increased funding was used to improve affordability), learning that lesson from the cases of Norway and Sweden.

In terms of the second possibility, of looking for contradictory cases with a single causal factor different from more successful countries, then the UK is perhaps the most important of the contradictory cases raised above as it is rated as the best overall performing system, but its combination of high government expenditure, and low voluntary health expenditure, means it appears as a contradictory case in the outcomes measure in which it scores poorly. Compared to Norway and Sweden (which also have that combination of funding factors, and alongside which the UK is often placed in typologies of health systems,

as we noted above), the UK is also a low overall spender on healthcare. This suggests that there is a route to improvement for the UK were to increase its overall health expenditure (and the gap to Norway is less than one percentage point as a proportion of total GDP). This analysis seems to be borne out by comparative studies within the different nations of the UK in terms of their health outcomes, which have gone different paths in terms of their organization post-devolution in 1999, but with similar funding mechanisms in place. This work suggest that the funding rises of the 2000s were the most significant factor in the improved health system outcomes in all nations in that decade (Connolly et al, 2010), and with the subsequent levelling off or even falling of those measures post-2010, after the impact of funding reductions has appeared, gives us strong reasons to believe that more could be achieved through greater funding of healthcare in the UK (Greener, 2018). Although the UK can be categorised as a 'public' health system in typologies of health systems, this more nuanced comparative exploration highlights the current and historically low levels of overall expenditure which make it different to Norway and Sweden, offering a criticism of UK health policy, but also a means to possible improvement for it as well.

In looking across the cases, there appear to be no consistently strong solutions for health systems based predominantly on high levels of voluntary health insurance regardless of the level of health expenditure in those economies. Australia is only a marginal exception to this as it is only just in the set of countries with high voluntary health insurance (0.51), and is more a member of the set of high out-of-pocket expenditure set (0.76). As such, Australia is best represented by the out-of-pocket expenditure fitness peak rather than the voluntary health insurance one – but this carries a strong recommendation that Australia does not increase its expenditure on voluntary health insurance beyond its current levels. Health systems utilising higher levels of voluntary health insurance do not appear to lead to good health outcomes, even in combination with other funding factors or different funding levels,

and if organising systems in this way is meant to lead to people taking greater responsibility for their health, then this is not appearing in measured health system outcomes. This has significant implications for countries such as the USA or Canada which appear to find themselves located within a fitness landscape point where high outcomes may not be achievable, but which face considerable interest mobilisation should they try and achieve change (Gordon, 2009). Indeed, the history of the health system in the USA points to the colossal problems involved in achieving any degree of change at all because of the fragmented political system, interest group mobilisation against change, and the difficulty of explaining change to the public when the system has grown so massively complex (Starr, 2013). It is hard to see how the USA can move to an alternative, but potentially higher, fitness landscape given these factors.

Turning to out-of-pocket expenditures, we see two fitness peaks, but one appears to lead to higher outcome and equity scores compared to the other. Australia links high out-of-pocket expenditure with relatively low GDP expenditure, and is a high performer in terms of overall ratings and outcomes, but falls short in terms of equity. Alternatively, high out-of-pocket expenditures can be combined with relatively high GDP and low voluntary health insurance for the solution present in Switzerland, Norway and Sweden. The major difference between these three countries is that the latter two have lower levels of out-of-pocket payments (although all falling in the set of higher scoring countries) and high levels of government expenditure compared to Switzerland.

As such, the potential for improvement for both Australia and Switzerland lies in increasing government expenditure to improve equity by removing the financial challenges it appears to create in health systems with high out-of-pocket expenditures, as well as with the strong recommendation that Australia does not increase its reliance on voluntary health insurance or out-of-pocket expenditure further. As the total health expenditure difference

between Norway and Australia is around one percentage point as a proportion of GDP, the gap is not huge, offering the potential for improvement for Australia to be attainable if government spending increases. This again offers us an example of the second type of fitness landscape improvement outlined above - of examining how a country can improve by looking at how single causal factors differ from other countries, and considering how it might be improved.

In terms of high government expenditure, a range of possibilities for achieving strong outcomes appear. High government spending combined with a low GDP health expenditure takes us to the case of the UK, which scores highly overall, and in terms of equity, but poorly in terms of health outcomes. If we combine high government expenditure with low voluntary expenditure then we have a solution that includes Norway and Sweden once again, but also Germany (poor outcomes), the Netherlands, New Zealand (poor equity) and the UK (poor outcomes). As such there is no consistent solution here for high government expenditure combined with high GDP, so achieving an improvement in the fitness landscape for each country is more contextual on which other causal factors are present. As such, utilising the language of fitness landscapes shows where improvements in health systems can be accomplished, but also helps explain why such changes may be difficult to achieve in practice because of the institutional inertia and established interest groups likely to be lobbying against such change (Wilsford, 1994).

The application of QCA allows us to construct a more nuanced typology of healthcare funding than systems based on public or private financing alone. Equally, the fitness landscapes we highlight also serve to problematise perhaps the most significant typology of welfare systems, that of Esping-Anderson (1990), in that they cast his concept of 'decommodification' (chapter 2) in a different light. Esping-Anderson defines decommodification in terms of the ability to which individuals or families can uphold a social

acceptable form of life without recourse to market exchange. However, as Powell and Barrientos (2011) make clear, this definition appears to suggest that citizens could opt out of all forms of work without any loss of income or welfare, and it would explicitly rule out health systems making even low use of either voluntary health insurance or out-of-pocket payments as both are mediated through market mechanisms. Although the findings above suggest that countries with high voluntary health insurance countries have all the low outcomes examined in this paper, and supporting a critique of the commodification of healthcare, some degree of out-of-pocket expenditure appears entirely compatible with high outcomes. The three countries in the solution set for achieving both strong equity and outcomes are all in the set of countries with high out-of-pocket expenditures, however, Norway and Sweden are only just in that set (with scores of 0.51 and 0.55), whereas Switzerland (0.95) appears to face greater affordability challenges than the other two countries (casting doubt on the Commonwealth equity calculation). It is also the case that Australia (0.76) is under-performing in relation to equity as well, despite its achievements on other measures. Medium levels of out-of-pocket expenditure, contra the Marmot quote above, appear to be compatible with achieving strong health equity and health outcomes, provided that government funding is also high (as it is in Norway and Sweden) – a surprising finding as we might have expected health equity to be more compromised by through such use of out-of-pocket expenditures.

If we use Esping-Anderson's (1990, p.52) the decommodification scores, then Australia, Canada, NZ, the USA and UK would fit within a 'liberal' regime, France, Germany and Switzerland as corporatist, and Netherlands, Norway and Sweden as social democratic. However, assessing countries according to their highest set membership would put the USA and Canada on a fitness landscape based on high voluntary health insurance (and poor outcomes), but with Australia on a landscape based around high out-of-pocket

expenditures and the UK, and NZ one based around high government expenditure. France and Germany too have their highest set score in one belonging to high government expenditure, but Switzerland clearly belong to the out-of-pocket landscape along with Australia. However, the Netherlands, Norway and Sweden all belong to fitness landscape of countries with high government expenditure. This presents us with a rather different clusterings to Esping-Anderson, but also one in which particular countries' position on different fitness landscapes depends not only on their highest-scoring funding source, but also both on their particular mix of other factors, as we have described above, and their 'fitness' in achieving the outcomes explored in the paper.

The approach taken in the paper also provides a language which can help us explore contradictory cases such as the UK, and how they might be further probed to consider how such systems can improve health systems outcomes. We should not expect health systems to conform neatly to our models – they are complex, path dependent, and include a range of factors no model could ever fully include. However, it is one of the strengths of QCA that it allows us, with good data, to dig deeper into individual cases, going, as Ragin advocated, back-and-forth between cases and the model to extract as much depth as possible from our analysis. Contradictory cases are a key means of achieving this, and are currently under-utilised in applications of QCA as a means of achieving greater insight from data.

REFERENCES

Abel-Smith B. (1994) *An Introduction to Health: Policy, Planning and Financing,* Basingstoke: Palgrave.

Berger M, Messer J. (2002). Public financing of health expenditures, insurance and health outcomes. *Applied Economics,* 34, 2105-13

Blank RH, Burau V and Kuhlmann E. (2018) *Comparative Health Policy,* London: Palgrave.

Carrera PM, Siemens KK and Bridges J. (2008) Health Care Financing in Germany: The
Case for Rethinking the Evolutionary Approach to Reforms. *Journal of Health
Politics, Policy & Law* 33: 979-1005.

Connolly S, Bevan G, Mays N. (2010). *Funding and performance of healthcare systems in
the four countries of the UK before and after devolution*, London: Nuffield Trust

Cremieux P, Ouelette P, Pilon C. (1999). Health Care Spending As Determinants of Health
Outcomes. *Health Economics,* 8, 627-39s

Deaton, A. (2013), *The Great Escape: Health, Wealth, and the Origins of Inequality*,
London: Princeton University Press

Dennett D. (1996) *Darwin's Dangerous Idea: Evolution and the Meaning of Life,* London:
Penguin.

Dusa A. (2018) *QCA with R: A Comprehensive Resource,* London: Springer.

Esping-Anderson, G. (1990), *The Three Worlds of Welfare Capitalism*, Cambridge: Polity
Press

Gordon C. (2009) *Dead on Arrival: The Politics of Health Care in Twentieth Century
America,* Princeton: Princeton University Press.

Greener I. (2018) Learning from New Labour's approach to the NHS. In: Needham C, Heins
E and Rees J (eds) *Social Policy Review 30.* Bristol: Bristol University Press.

Hudson J and Kuhner S. (2010) Beyond the Dependent Variable Problem: The
Methodological Challenges of Capturing Productive and Protective Dimensions of
Social Policy. *Social Policy and Society* 9: 167-`179.

Jessop B. (2002) *The Future of the Capitalist State,* Cambridge: Polity Press.

Johnson J, Stoskopf C and Shi L. (2017) *Comparative Health Systems,* Burlington: Jones and
Bartlett Publishers.

Kaminska M and Wulfgramm M. (2018) Universal or commodified healthcare? Linking out-of-pocket payments to income-related inequalities in unmet health needs in Europe. *Journal of European Social Policy* online early: 1-16.

Kingdon J. (2011) *Agendas, Alternatives and Public Policies (updated second edition),* Boston: Longman.

Lawson N. (1991) *The View from No. 11,* London: Corgi.

Marmot T and Wendt C. (2012) Conceptual frameworks for comparing healthcare politics and policy. *Health Policy* 107: 11-20

Marmot M. (2007). Achieving health equity: from root causes to fair outcomes. *Lancet,* 370, 1153-63

Marmot M. (2015). *The Health Gap: The Challenge of an Unequal World*, London: Bloomsbury

Mischen P and Jackson S. (2008) Connecting the Dots: Applying Complexity Theory, Knowledge Management and Social Network Analysis to Policy Implementation. *Public Administration Quarterly* 32: 314-338.

Pierson P. (2004). *Politics in Time: History, Institutions and Social Analysis*, Woodstock: Princeton University Press

Powell, M. and Barrientos, A. (2011), An Audit of the Welfare Modelling Business, *Social Policy and Administration*, 45,1,69-84

Ragin C. (2008) *Redesigning Social Inquiry: Fuzzy Sets and Beyond,* Chicago: University of Chicago Press.

Ragin C. (2014) *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies,* Oakland: University of California Press.

Schieber G, Poullier J-P and Greenwald L. (1991) Health Care Systems in Twenty-Four Countries. *Health Affairs*: 22-38.

Schneider C and Wagemann C. (2012) *Set-Theoretical Methods for the Social Sciences,* Cambridge: Cambridge University Press.

Schneider E, Sarnak D, Squires D, et al. (2017) *Mirror Mirror 2017* New York: Commonwealth Fund.

Smith P, Stepan A, Valdmanis V, et al. (1997) Principal-agent problems in health care systems: an international perspective. *Health Policy* 41: 37-60.

Starr P. (2013). *Remedy and Reaction: The Peculiar American Struggle Over Health Care Reform*, London: Yale University Press

Vis B. (2009) Governments and unpopular social policy reform: Biting the bullet or steering clear? *European Journal of Political Research* 48: 31-57.

Wendt C. (2015) Healthcare Policy and Finance. In: Kuhlmann E, Blank RH, Bourgeault IL, et al. (eds) *The Palgrave International Handbook of Healthcare Policy and Governance.* London: Palgrave, 54-68.

Wendt C, Firisina L and Rothgang H. (2009) Healthcare System Types: A Conceptual Framework for Comparison. *Social Policy and Administration* 43: 70-90.

Wilsford, D. (1994), Path Dependency, or Why History Makes It Difficult but Not Impossible to Reform Health Care Systems in a Big Way, *Journal of Public Policy*, 14, 3, 251-283