

[Joan Costa Font](#), Joan Rovira Forns, Azusa Sato
**Participatory health system priority setting:
evidence from a budget experiment**

**Article (Accepted version)
(Refereed)**

Original citation:

Costa-Font, Joan, Forns , Joan Rovira and Sato, Azusa (2015) *Participatory health system priority setting: evidence from a budget experiment*. [Social Science & Medicine](#). ISSN 0277-9536

DOI: [10.1016/j.socscimed.2015.10.042](https://doi.org/10.1016/j.socscimed.2015.10.042)

Reuse of this item is permitted through licensing under the Creative Commons:

© 2015 The Authors

CC-BY-NC-ND

This version available at: <http://eprints.lse.ac.uk/64144/>

Available in LSE Research Online: October 2015

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

SSM-D-15-00563R4

Participatory Health System Priority Setting: Evidence from a Budget Experiment

Joan Costa Font,
London School of Economics and Political Science

Joan Rovira Forns
University of Barcelona

Azusa Sato
London School of Economics and Political Science

Contact Address: Dr. Joan Costa-Font, London School of Economics, Cowdray House,
Houghton Street, WC2A 2AE. E-Mail: J.Costa-Font@lse.ac.uk

Abstract

Budget experiments can provide additional guidance to health system reform requiring the identification of a subset of programs and services that accrue the highest social value to ‘communities’. Such experiments simulate a realistic budget resource allocation assessment among competitive programs, and position citizens as decision makers responsible for making ‘collective sacrifices’. This paper explores the use of a participatory budget experiment (with 88 participants clustered in social groups) to model public health care reform, drawing from a set of realistic scenarios for potential health care users. We measure preferences by employing a contingent ranking alongside a budget allocation exercise (termed ‘willingness to assign’) before and after program cost information is revealed. Evidence suggests that the budget experiment method tested is cognitively feasible, incentive compatible. The main downside is the existence of ex-ante “cost estimation” bias. Additionally, we find that participants appeared to underestimate the net social gain of redistributive programs. Relative social value estimates can serve as a guide to aid priority setting at a health system level.

Keywords: Catalonia, budget pie, budget experiments, willingness to assign, preference revelation, , priority setting.

1. Introduction

Health care decision making under publicly financed systems is a ‘community commodity’. In other words, rather than a market allocation, health care decisions in such systems are the result of a ‘socio-political market’, whereby social exchanges highly depend on social norms (e.g., equal access for equal need) (Heyman and Ariely , 2004) and political processes. Nonetheless, although decisions in publicly funded health systems are generally grounded in a political mandate, rules of democracy alone do not guarantee that the population’s values are accounted for, especially the socially excluded population who are less likely to be electorally relevant. At best, traditional decision-making imperfectly reflects broad preferences. Relying on health maximisation is equally problematic. For instance, Lindholm *et al* (1997), using interviews of Swedish politicians, reveal that public preferences do not follow cost-effectiveness criteria in the presence of inequality. Stafinski *et al* (2011) also highlight the importance of social values and distributive preferences and other values (personal responsibility, caregiving responsibilities etc) in resource allocation decisions.

Under such a scenario, decision-makers’ preferences for health programs are arguably not in line with public preference, which calls for methods to analyse public participation in health system decision-making that supplements cost-effectiveness analysis. Whitty and Littlejohns (2014) find that in Australia, health decision making competes with several inter-related social values, and therefore suggest that participation can be a way to enhance the process. However, when multiple goals are at stake, decision making must

be contextualised through Multi-Criteria Decision Analysis (MCDA) where several outcomes - including wider social effects of an intervention - are accounted for (Hauk et al, 2003, Mitton et al, 2011), and macro level health system approaches such as the so called 'program budgeting and marginal analysis' (PBMA) as per Mitton *et al* (2003) are used. Methodologies which attempt to improve the existing elicitation of preferences for health programs, (such as direct elicitation willingness to pay exercises) and discrete choice experiments, present one way forward. However, as we discuss below, willingness to pay studies do not show consistent valuations.

In any event, when such exercises are undertaken, preferences over a given set of health programs may exhibit at least two different dimensions: an individual (selfish) and a collective (altruistic) dimension as described by Margolis (1982). Of course, the share of one value or the other may be cultural specific. Hence, the individual intensity of preferences for different health programs is interdependent or interpersonal (Mooney 1994; Bazerman et al, 1992).

Both laboratory and natural experiments call for a redefinition of preference elicitation mechanisms, so as to pick up individuals' collective sacrifice (rather than exclusively rely on individual welfare gain). Specifically, we propose a budget pie experiment as a straightforward way to elicit preference in settings where spending cuts (or expansions) to collectively funded health services take place, in order to add an additional dose of realism. The budget pie experiment focuses on soliciting responses to allocate a fixed budget to a set of potential alternatives (programs), which are sometimes referred to as

constant sum comparisons (or scaling), and can be implemented in pairs of programs or a limited number of programs. The advantage of budget pies is that they make the respondent consider the whole budget, and hence the sacrifice of allocating more funds to one program over another. As a result traditional 'yea saying' behaviour of willingness to pay experiments is reduced (Olsen and Donaldson, 1998).

Budget pie experiments have been examined in the health context in Skedgel *et al* (2013) in a choice context. Indeed, Skedgel *et al* (2013) draws from discrete choice techniques and budget pie strategies with varying health system attributes. They find that budget pie techniques are more aligned with societal preferences, but discrete choice respondents appear to be more consistent. Among the dimensions of benefit examined, health gain was the most valued attribute. Schwappach (2003) finds that by using a budget pie allocation mechanism, only a small fraction of individuals prioritised health gain, with many willing to sacrifice some health gain to attain more equal distribution of resources.

Alternative willingness to pay (WTP) approaches tend to focus on individualistic decision settings to measure the social value of a program, which unsurprisingly does not necessarily give rise to consistent outcomes (Olsen and Donaldson, 1998), and rarely contemplate explicitly the collective budget restrictions that individuals face as 'citizens' in addition to that of individuals or potential health care 'users' (Eckerlund, *et al*, 1995). In contrast, a 'fair share allocation rule', as proposed by Margolis (1982), takes into account the observation that when humans allocate resources they always make a trade-off between individual and social values. Such ideas have been recently corroborated

through evidence of motivational crowding out (Frey and Jegen, 2001, Heyman and Ariely, 2004). Kahneman and Ritov (1995) find that in a collective context, WTP measures may elicit the willingness to contribute, rather than consumer preferences.

In this paper, we report on a method for measuring collective sacrifice with something more than an individual's utility (lathering preferences) by relying on budget allocation exercises with and without the disclosure of program cost information. While some research has shown that it is possible to elicit the joint WTP for multiple programmes (Protière *et al*, 2004; Robins and Simonsen (2002)), results are inconsistent. Furthermore, although alternative non-monetary elicitation techniques such as willingness to wait (Thomas *et al*, 2000) have been put forward, they do not necessarily allow for generalisation to multiple programmes. An alternative is a structured value referendum (McDaniels, 1996). Nonetheless, when budget information is provided for only one service, but not put into context with the costs and trade-offs related to other programs, the very mention of a tax, all else being equal, will give rise to less support for that service. Hence, budget experiments can capture a wide collective decision making dimension when other methods fail (Bergquist et al, 2004; Petters, 1996). Rather than focusing on eliciting the marginal willingness to pay for consumption, trade-offs are between various programs, by allocating budget increments (or cuts) to the programs. If preference elicitation methods are to guide priority-setting decisions, it is necessary to examine other preference elicitation methods, which deal with sets of multiple health system benefits, and collective decision-making scenarios, which take into account opportunity costs of funding other programs. For instance, Franken and Koolman (2013

page 29) distinguish between utility derived from interaction between people and the health system (“how people are treated by the health system”), and utility obtained from living under institutions (“how allocative and redistributive decisions are taken”). In addition to process utility (or benefits associated with the delivery of care process that might not improve health outcomes), other important benefits are those associated with equity concerns. Equity in the financing of health care is health system benefit in the ranking of health systems (Murray et al, 2000). However, there is some criticism on weighting methods and the implicit assumption of a common ideology (Richardson, 2003) or the same welfare trade-offs across countries (Smith, 2002).

We agree with Franken and Koolman (2013) that the structure of a health system depends in large part on society’s choices concerning resource allocation; and hence it is important to elicit ex ante socially inclusive personal preferences in social justice valuations. However, it is unlikely that individuals will have formed preferences for concepts unless there has been a deliberative process before and health decisions may reflect social choices. Hence, we have developed a budget experiment instead which includes group deliberation prior to individual choices being made. From a social choice perspective, participants were not told, and could not guess, how preferences were going to be aggregated and hence we would not expect any strategic behaviour. Second, as preferences were elicited by groups, there were no advantages from free riding (as we show later, we could not find any individual characteristic correlating with valuations).

In the following, we describe the subject background, the need for using budget experiments as an alternative to willingness to pay studies and we propose a specific budget experiment based on a realistic health care reform and examine the evidence to inform both the applicability of the methodology as well as the results of the experiment itself. That is, we aim to elicit what we regard as the ‘collective sacrifice’ undertaking a budget experiment that simulates a political market, where individuals are asked to participate in realistically setting their health system reform budget before and after program costs are revealed. This equates to an alternative measure of value to individual sacrifice.

To preview our results briefly, we find that budget experiments under a participatory setting did not receive significant protest responses and allowed revealing a consistent prioritisation of health care programs, although we find evidence of some ‘ex-ante’ cost estimation. The valuation of different programs are indicative that programs which were lowly priced, when costs were revealed, improved their final ranking, whereas programs where cost was high, reduced their final ranking. From these results we conclude that individuals seem to be cost sensitive although the magnitude might not be large enough to change the average order of the program rankings. Our findings seem to suggest that individuals perceive more than health gain in health systems reform but importantly results suggest that in allocating health care funds, process utility benefits do not garner a high budget, and programs improving access to health care coverage, although highly valued, would not be sufficiently funded.

2. Background

2.1 Eliciting the social sacrifice for health care programs

Governmental decisions in areas where the potential impact of social welfare is high may suffer from information deficiencies (Willoughby and Melkers 1998). Besides from making decisions based on technically determined results, it is possible to determine ex-ante, through experimental designs, on how best to distribute the existing budget. This implies eliciting preferences for different program budgeting options, which to an extent removes political pressure from the estimation of budgets for each specific program. Contingent valuation studies suffer from the existence of order effects when multiple goods are subject to valuation (Payne et al, 2000).

It is widely acknowledged that different elicitation methods give rise to different preference rankings. In health, programs may be explicitly ranked, or estimations of willingness to pay extracted. In the former, a marginal approach is often taken – for example, programmes for the expansion of health care coverage are anchored to the most preferred option (Olsen *et al*, 2005). However, evidence suggests that the value of public goods falls when the good is valued with a sequence of other goods (Carson et al, 1998; Carson, 2000). This feature cannot be explained by economic theory unless we accept that individuals must be confronted with realistic decision-making contexts where real opportunity costs are taken into account. One of the alternatives being proposed is to

introduce direct comparison of programs. Indeed, while some individuals support certain programs, others oppose them strongly, and this is likely to be linked to distribution implications of some programs (Corneo and Fong, 2006).

In political and social markets, individuals may be modelled as following ‘expressive behaviour’, where they are concerned about the social, rather than individual, impact of health programs. Kahneman and Ritov (1997) find that in a public goods context, WTP measures elicit willingness to contribute rather than consumer preferences per se.

Another problem in valuing health care is the presence of some moral satisfaction as health care investment is seen as a good cause (Kahneman and Knetsch, 1992).

Donaldson et al (2002) find that consistency of WTP studies depends on the alternative courses of action under comparison. This might indicate that an *individual's reference point* in valuing goods differs when different programs are valued together, and in particular, specific program characteristics might stimulate certain utility dimensions compared to others (Tversky and Simonson, 1993). Furthermore, Donaldson et al (2002) find that respondents do not discriminate between different treatments when they are unaware that alternatives exist, which indicates that preferences are sensitive to the fact that different courses of action are made explicit to decision makers. However, Olsen et al (2005), using data from different countries, find little consistency across explicit rankings and willingness to pay estimates. This shows that that WTP estimates are not particularly appropriate to elicit relative values of competing health care programs. In other words, to judge the suitability of a methodology to elicit preferences to guide health care reforms one would expect the implicit ranking revealed by the value measure to

correlate highly with the explicit ranking individuals provide, so long as elicitation measures meet convergent validity criteria. Finally, monetary preferences are argued to be subject to monetary illusion, so that individuals would tend to reason on nominal, rather than real, terms (Shafir *et al*, 1997).

Another reason for measuring collective sacrifice is because individual valuations reflect some form of social valuation, and hence different elicitation methods might trigger motivational crowding out (Bowles, 1998). Frey and Oberholzer (1997) argue that the introduction of monetary payments might reduce the intrinsic motivation to behave altruistically or perform civic duty. Similarly, individual revealed preferences from choices both real (states preferences) and hypothetical (contingent valuation experiments) might differ from true normative preferences because social constraints can trigger different protest responses, as well as a variety of well-established biases such as limited experience and passive choice, complexity, inter-temporal dimension and third party marketing (Beshears *et al*, 2008).

2.2 Participatory budget experiments

There have been some experiences of participatory budgetary priority setting exercises in the US, in defining drug coverage (Sacramento Health Care Decisions, 2006). There is growing interest in involving citizens in policy development to ensure that decisions are legitimate, and reflect broad social values (Bombard *et al*, 2011). There is growing interest in involving a broader range of stakeholders in public policy development and

HTA-related policy more specifically including patients and service users (or 'consumers'). This enthusiasm stems from a need to ensure that decisions are informed, transparent and legitimate (Abelson, 2003, 2007). Universal access, 'choice', and 'quality care' were identified as core values that should be considered in the evaluation of health technologies and ensuing recommendations.

Based on these experiences, preferences are best elicited using collective decision making settings, so we propose a budget pie experiment format termed 'willingness to assign' (WTAS) method, which will expand the traditional willingness to pay (WTP) format to a collective setting. As in a traditional budget pie experiment, participants were asked to act as collective decision makers, and allocate the maximum amount they would wish to assign to a set of given health care programs under a fixed budget constraint (funded by their own taxes). They could reassign their allocations so as to elicit marginal valuations for a program coverage expansion where tradeoffs between health care programs were made explicit. The interpretation of results applying this method has been carried out elsewhere (Costa-Font and Rovira, 2005). The experiment was undertaken by grouping individuals of similar social class and skills to enable the examination of differences in preferences across groups. We employed a deliberative approach, given that it is more likely to elicit informed judgements when it provides a space for the exchange of experiences and opinions (Yankelovich, 1991). Finally, a direct ranking of programs was asked for, to test the consistency of individual responses. The method is similar to that of Blomquist et al (2004), but the difference lies in that we rely on a deliberative process

rather than a survey, and we attempt to elicit the relative trade-offs between programs within the health sector rather than allocation between sectors.

3. Methods

3.1 Experimental Scheme

The experimental scheme can be contemplated in terms of a three-stage process:

In the first step ('benefit identification' exercise to elicit health system benefits), a set of focus groups aimed at benefit identification. Each focus group was held at a different geographical location within Catalan provinces. Moderators ensured that every participant contributed to the discussion with opinions and arguments (to avoid more assertive participants to manipulate the course of the debate) and discussion were recorded and transcribed.

In the second stage (the 'benefit assessment' exercise which was only partially participatory and was mainly done following expert assessments), a set of ten hypothetical health care programs were developed, each of which would be expected to be able to deliver the benefit improvements identified in the first step. The concept of health system benefit was defined broadly as any welfare increase resulting from a new health care programme. The programs were designed (with consent from participants) by officials of the Catalan health system to make them realistic interventions that met the expected benefits, which the health system at the time was not offering.

In the final, third stage (program valuation stage), individuals were asked to take part in a benefit ranking exercise (by order of preference), a budget pie (also referred as WTAS) exercise and a cost based ranking exercise of the programs designed in the previous stage. This last section began with an explanation of the program and a brief discussion (lasting 40 minutes) followed by a contingent ranking exercise (30 minutes) of the programs without preliminary knowledge of its cost. Then, a budget pie (or willingness to assign WTAS) exercise and in some cases a willingness to pay taxes question was answered in 20 minutes and results were discussed in under 10 minutes. Finally, a rank ordering exercise was repeated but now the program costs were revealed to participants without any further discussion. In total, the exercises took about two hours until completion.

This stage included a survey by using cards where every program to be valued was briefly described and participants were rewarded in each group. Before every session, participants were requested to fill in a basic personal questionnaire. Both the documentation and the questionnaires alongside a reference guide were mailed beforehand.

The implementation of the research project benefited from both the participation and advice of two social researchers (different from the health system experts) who helped to identify the characteristics of the representative social groups of study. Their aim was to identify groups that brought together a mix of people who shared similar backgrounds. Participants were informed about the purpose of the exercises as well as basic details of the Catalan health system in an accessible way. Ethical clearance was obtained in

October 1999 from University of Barcelona social and economics science ethics committee, and data collection took place between 1999 and 2002.

3.2 Participant selection

When a large set of health programs are to be prioritised, standard surveys present limitations. Small-scale experiments, and particularly a participatory group experiment format including the focus group approach, can allow for more in-depth investigation of the association between individual priorities and values, reduce misunderstandings of questions where advantages of each programme can be discussed and adequately capture the decision-making problem. We recorded information about age, education health care interventions, residence and employment history which helped further group stratification in groups ranging between 6 and 15 participants (including the pilot experiment). Table A1 describes the profile of the participants in each group and the number of participants. Altogether they make up a sample of 88 observations clustered in different participant groups.

3.3 Elicitation instruments

Different methods show clearly their different aims (reflected in Table 3). The first method allows one to check consistency with WTAS (or budget pie allocations) results, since ordinal preferences should show a significant association with cardinal measures.. The third cost priority question tries to provide some notion of how consistent the

individual responses were. Since the program cost is a negative dimension (even under collectively funded systems as it reduced funding for other programs), we expect it would trigger a reduction in the level of priority assigned to every program. Finally, we took into consideration an alternative and more traditional elicitation technique for comparison, eliciting the WTP taxes for each health program. Consistency of respondents' valuations was measured using Spearman correlations, namely correlation coefficients of ordinal outcomes. The assumption is that observed ranking can be expressed as follows: $R = V + e, E(e) = 0, Var(e_{ij}) = 0$, where the ranking of a group can be understood as an addition of a "real" ranking plus a random error, with zero mean and constant variance. The null hypothesis is that there is no correlation between the rankings and random errors. From this assumption the Spearman coefficient is: $\sigma^2_{xv} = \frac{\sigma_v^2}{\sigma_x^2}$, varying between -1 and 1.

The different elicitation tools were the following:

- A. **Contingent ranking**: this method hypothetically captures the ordinal priority of different programmes. Participants provide their rankings on a scale of 1–10.

Question (1): *Please could you rank each programme (after every single programme had been explained and comprehension was ensured) from 1 (the least preferred) to 10 (the most preferred)?*

B. Budget pie or willingness to assign (WTAS). This was based on an open-ended valuation question that considered a hypothetical budgetary increase (that does not exceed the total budget). The purpose is to elicit the maximum amount individuals would assign to a set of given health care programs under a fixed budget constraint. The idea underlying this method is one of identifying relative values and trade-offs, therefore any budget size variation according to the monotonicity criteria should not affect the preferences elicited, as they were considered in terms of relative values.

Question 2: *Imagine being asked to allocate the health care budget as if you were the decision maker. You have 1,000 million € budget to assign the money to each health programme in the system. How would you, as a 'citizen', assign public resources of a realistic list of health programs that have been designed from a previous list of health system benefits you contributed to reveal?*

The advantage of the budget pie (WTAS) format employed here is that it allows us to compare the valuation of different programs using a similar measure, which in turn makes it easy to identify (partial) trade-offs between different programs that satisfy heterogeneous benefits. We consider the decision making unit as small groups (n_i), representative of a societal profile ($N = \{\sum n_i\}$). Let us assume a hypothetical health system resource allocation (e.g., Catalan Health Service (CHS)) and assume planned budget increase (e.g., 240 million €), to fund a set of random number (e.g., ten) of "new programs". The total cost of those (ten) programs is in excess of the fixed capital budget;

therefore some programs may not be funded. Hence, the experiment implicitly attempts to assist rationing decisions in a multidimensional benefit or utility framework. The money metric nature of budget pies can be of some additional use if we are interested in a cost-benefit decision making framework. The context of policy decision-making was explained to respondents as stating that the government was considering a coverage expansion that would satisfy the potential health system benefits revealed in a prior exercise. The list of programs was carefully explained to participants, a written list was posted in a white board, and a set of programs cards describing each program features. Participants were asked to act as if they were advisors of the Catalan Health Plan reform board. To further enhance a realistic scenario, a summarised (and again realistic) public health system budget was revealed and participants were made aware that every program has been partially paid by their taxes. The latter was expected to enhance the notion of ‘collective sacrifice’ which we expect the elicitation technique to identify. Finally, it is important to note that the allocation exercise requires participants to allocate funds in addition to the existing budget at the time of the study, which was announced and explained to participants.

One of the potential problems is the potential complexity and limited experience of participants as decision makers, rather than relying on survey based methods. In contrast, Blomquist et al, (2004) and Skedgel et al (2013) we use a used a participatory and group allocation mechanism. Participants could then engage in a discussion as a group, on the reasons for their collective decisions. Given that not all groups were equally dynamic and participatory, in some cases extra effort was required to engage participants and aid

responses; for example, some people were reminded to think as if they had to purchase food for their family for which they had a fixed amount of money. Moreover, they were not required to expend the whole budget; any left over money would be assigned to other public sector benefits.

C. Cost based Ranking: To understand the consistency of WTAS responses and their sensitivity to the cost of every program the previous program ranking was repeated but this time incorporated information on the health care program costs. Indeed, the hypothetical cost amount was calculated according to whole programs costs (by senior officials of the Catalan health service to make is realistic as possible). The design was such that the total set of programs should not be completely covered by the budgetary increase. Hence, it would force individuals to make choices based on a notion of ‘collective sacrifice’. The program cost may be conceptualized as a “negative benefit” to be weighed against the benefits each program attained. In other words, participants would prefer a program that involves a reduced cost to an equivalent program – yielding the same benefits involving higher costs. Therefore, a higher degree of cost sensitivity might be interpreted as economic rationality. The latter was tested by comparing cost based ranking (after the program costs were revealed), with both the implicit ranking from budget allocations and contingent rankings of programs (before program cost information was revealed to participants). The exact cost based ranking question is described below:

Question (3): Please could you rank every program (after every program was explained and understood by respondents) from 1 (the least preferred) to 10 (the most preferred) after the costs associated with each of the following are revealed?

Finally, a willingness to pay taxes question, using a payment card mechanism for each program, was included for comparison. This question as described below in Table 3 was elicited by asking: ‘how much extra tax would you be willing to pay annually to have this program covered by the (Catalan) health service? (A payment card including N/A option was included with quantities ranging from 0 to 500€ or more).

4. Results

4.1 Benefit identification and analysis

The identification of health system benefits tries to reflect, in a simplified way, basic outcomes. The outcomes that every participant reveals can be positive, identifying benefits for the health care system, or negative, highlighting shortcomings of health care systems. The former is represented by a positive (+) mark (when participants felt the health services ought to deliver) and health system deficiencies are marked with a negative (-) mark (when the participant felt that the health service was failing to deliver something they had expected to be delivered). It is important to note that researchers did not predetermine nor define the benefit dimensions, but respondents, in their social groups, reached a consensus on them. After the identification of health system outcomes, the first part of the experiment is completed (see Table 1). Importantly, this set of exercises revealed that people identify different benefits accruing from the health system, among which health gain is the most acknowledged.

[Insert Table 1 about here]

4.2 Health program identification

Given that health system benefits are usually tailored by using sets of programs and policies, we converted all choices into an available set of alternatives from which they can be valued and identified. A set of ten programs was designed and benefits were classified in five dimensions: coverage, accessibility, quality, public health and social. With the assistance of some health system officials, we estimated an approximate but realistic cost for each program (at the time of study) and made sure that the proposal would increase status quo health system coverage so that it remained a realistic exercise.

[Insert Table 2 about here]

4.3 Preference elicitation

Table 3 describes the different elicitation techniques employed in the exercise, its objective and limitations to be tested. We began with a contingent ranking of existing health programs and asked participants to rank them on a scale of 1-10. This elicitation exercise attempted to capture the ordinal preferences of participants for every program.

After individual elicitation we published the results and a discussion took place about the reasons for valuation. Table 4 suggests that when we examine rankings, the population reveals a priority for extending coverage of dental care and co-payment reduction (programs 9 and 10), both programs that improve equity in the access to health care, followed by program 8, which aims to improve information for prevention purposes. In contrast programs that target organisational aspects of the NHS receive lower rankings, suggesting that individuals might not be able to fully appreciate the value of managerial practices and organisation structures. Standard errors, reflecting the variability or consensus on the estimates, were similar except for the highest valued programs where we observed half the variability in the valuation.

[Insert Table 3 and 4 about here]

As a second preference revelation exercise we elicited cardinal and implicit trade-offs between health care programs by allocating an expected budget which is collectively funded. Consistently, the highest and the lowest programs average funding assigned correspond to the equivalent rankings as explained before. Only one program, medical checkups, did not receive comparable funding to programs 4 and 5 despite receiving a higher rank. As we argue in the conclusion, we interpret this finding as suggestive that participants were taking the implicit cost into consideration when formulating their value assessments.

Since our data are ordinal the best way to test for validity is to look at the consistency of rankings. Consistency estimates are reflected in Table 4, in which the Spearman correlation coefficient is always consistently positive and in many cases significant with values around 0.4-0.7. Compared willingness to pay taxes, we did not find a sign reversion in WTAS responses. This could be because we did not employ the standard referendum approach and instead used a payment card. Another explanation is that in a collectively funded health system, health services are often used irrespectively of its cost. Finally, when we compare WTAS values and costs for each program one can notice that although the two programs bringing equity benefits are highly prioritised and valued, equity programs are unlikely to be funded given their high costs. This result leads us to the final exercise of cost sensitivity. Table 5 displays more information on cost ranking *vis-a-vis* an implicit preference ranking from budget pie (WTAS) allocations as well as cost rankings using the non-parametric test of the equality of rankings. Importantly, results indicate that although the rankings are very similar in average values they fail to reject the null hypothesis of no equality for all programs but number 2. However, there is some level of sensitivity to the cost as Table 5 reveals.

Upon revealing the cost of programs, participants were asked to rank the programs again (which we refer to as 'cost ranking'). As Table 5 suggests we did not find significant changes in aggregate ordering, which could suggest that differences in cost may not be perceived to be large enough to overcome differences in benefits. Another explanation lies in that individuals might uphold some preliminary estimate of the cost of certain programs, and hence their valuations might be biased (ex-ante cost-estimate bias).

Alternatively, it may be also due to lexicographic order, where differences do exist but they do not manage to change the order of preference. If there are changes in the ordering, then we would be capturing how responses vary in their valuations and preferences when cost information is provided.

[Insert Table 5 about here]

4.4 Consistency and interpretation

Results suggest relevant differences in terms of the consistency criteria across groups. We estimated the distance between the two different rankings weighted by the maximum possible difference, e.g., if three programs are compared and rankings range from 1 to 3, the maximum difference is 2. We find that although participants identified ‘process utility’ as an important health system benefit, when they are asked to assign value health care programs dealing with such benefit dimension, they obtain a comparatively low priority value. One explanation may be due to the programs selected. Programs dealing with cancer may be overvalued, with people showing a moral satisfaction effect; however this is not reflected in check-ups. Although respondents do not appear to assign a large value to process utility, they do value programs leading to equity benefits (see Table 5). Moreover, moral satisfaction may fail to express itself when more than one program is valued. The reason is that individuals in this experiment are aware that the assignment of an excessive monetary valuation to one program may reduce the available resources for alternative programs.

5. Discussion

This study has sought to examine the feasibility of a participatory budget experiment to identify and value preferences for competing health care programmes in a collective funding health system. We have argued that unlike contingent valuation methodologies, a budget experiment appears a priori to be more in line with the sort of health system decision making in collectively financed health systems. However, the evidence available so far on the feasibility of budget experiments is very limited. In addition, we have relied on a participatory methodology allows individuals to identify the benefits and then convert such benefits into programs, to then value them in a way that suits their social preferences and accounts for public sector budget constraints. This is consistent with Scuffham et al (2010) who draw upon a discrete choice model to examine health system preferences in the UK and Australia. However, our experiment employs instead the enormous potential of deliberative methods (Mitton *et al* 2009), which are still underdeveloped and worthy of further exploration particularly in the context of public insurance coverage expansion.

The contribution of the paper lies in testing a participatory preference elicitation methodology and an elicitation technique that does not exhibit protest responses and allows the identification of a 'collective sacrifice'. Furthermore, it overcomes some of the problems identified in other budget pie studies that rely on survey data (Skedgel *et al*, 2013). The method combines a set of preference elicitation tools that is of policy use

when making resource allocation decisions in the context of health care, especially at times of budget cuts, but can which can also be implemented across the board at all times. We have found that individuals value more than health gain but programs attaining mainly process utility benefits do not garner a high allocation. In contrast, we find that programs improving access to health care coverage, although highly valued, would not be sufficiently funded in the context of the experiment. This result is suggestive of a quality of budget experiments, namely that by valuing more than one program at a time, the collective opportunity costs (e.g., not funding other health programs) might have curtailed the potential for moral satisfaction in the expression of monetary valuations. Finally, we find from a cost based ranking exercise that participants seem to be cost sensitive although the magnitude might not be large enough to change the average order of the program rankings. Finally, we have found no evidence of an overall lexicographic order at both group and individual levels.

Limitations

The proposed methodology has some important limitations: First, respondents might already have an estimate about the costs of different programs, and hence suffer from an ex-ante “cost estimation” bias. Indeed, we find that individuals, in reporting their WTAS estimates, might take implicit costs into consideration (ex-ante cost estimation bias). Second, in taking cost into consideration, they appear to underestimate the amount of funding required to reimburse a highly valued programs that extend health care converge and increase equity of the health system. Third, another of the main problems is the existence of fiscal illusion by individuals that participated in the experiment (Pluviani,

1972). That is a certain inability to elicit the monetary value of a good accounting for purchasing power changes over time. Fourth, a more general weakness of public participation mechanism in collective decision making lies in dealing with individual's reluctance to participate which in turn over- represents the preferences of certain population groups. Finally, another potential issue to examine in future work is how to include a preference for public savings or investment outside health care which exert influence on the health system.

6. Conclusion

Health care reforms in publicly funded health systems calls for further reliance on citizen's valuation to guide priority setting. Most available preference elicitation techniques that health economists use either rely on measures of health gain alone (e.g., Cost-Effectiveness Analyses), or on individualistic (market based) valuations which conflicts with the collective nature of health care decision making (e.g., willingness to pay directly or through taxes). This paper has examined the evidence of a participatory budget experiment (budget pie) to guide health care decision making. On the basis of this experience we conclude that budget experiments are an operationally feasible way for citizens to face the complexities of real decision-making, which are subject to significant moral and preference trade-offs. Two main limitation of budget experiments are identified, namely the presence of some ex-ante cost estimation, and participants appeared to underestimate the net social gain of redistributive programs. Further, some research shows that the views of participants in deliberative processes change after being

involved in such experiences (Stafinski et al 2014). Such areas are potential future research topics.

References

- Abelson, J., Forest, P. G., Eyles, J., Smith, P., Martin, E., & Gauvin, F. P. (2003). Deliberations about deliberative methods: issues in the design and evaluation of public participation processes. *Social Science & Medicine*, 57(2), 239-251.
- Abelson, J., Giacomini, M., Lehoux, P., & Gauvin, F. P. (2007). Bringing 'the public' into health technology assessment and coverage policy decisions: from principles to practice. *Health Policy*, 82(1), 37-50.
- Abelson, J, Gerlier Forest, P, Eyles, J, Casebeer, A, Mackean, G and the Effective Public Consultation Team. Will it make a difference if I show up and share? A citizen's perspective on improving involvement processes for health system decision-making. *J Health Service research and Policy*, 9: 205-212.
- Arrow, K.J (1951). Social Choice and Individual values. (Cowles Commission Monograph, n°12) New York: Wiley.
- Baserman, M, Lowenstein, G and White, SB (1992). Reversals of Preferences in Allocation Decisions: Judging an Alternative v Judging Among Alternatives. *Administrative Science Quarterly*, 37, 220-240.
- Beshears, J & Choi, J J., Laibson, D & Madrian, B C (2008). How are preferences revealed?. *Journal of Public Economics*, 92(8-9):1787-1794.
- Blumstein (1997). The Oregon Experiment: The Role of Cost-Benefit Analysis in the allocation of Medicaid funds, *Social Science and Medicine*, 45 (4): 545-554.
- Bombard Y, J Abelson D. Simeonov, F.P Gauvin (2011). Eliciting ethical and social values in health technology assessment: A participatory approach. *Social Science & Medicine* 73 (2011) 135- 144
- Blomquist, GC M A. Newsome, and D. B Stone (2004). Public Preferences for Program Tradeoffs: Community Values for Budget Priorities. *Public Budgeting & Finance* 49-51
- Boyle, KJ, Welsh, MP and Bishop, RC (1993). The role of question order and respondent experience in contingent valuation studies. *Journal of Environmental Economics and Management*, 25: 80-99.
- Calnan, M (1997). Citizens, users and health care. *European Journal of Public Health*, 7:1-2.
- Carson, RT, Flores, NE, Hanemann, WN (1998). Sequencing and valuing public goods. *J Environmental Economics an Management*, 36: 314-323.
- Carson, RT (2000). Contingent Valuation: a user's guide. *Environmental Science Technology*, 34: 1413-1418.

- Clark, TM (1974). Can you buy a budget pie? *Policy and Politics*, 3: 3-31.
- Costa-Font, J and Rovira, J (2005) Eliciting preferences for collectively financed health programmes: the 'willingness to assign' approach *Applied economics*, 37 (14). 1571-1583
- Cookson, R and Dolan, P (1999). Public views on health care rationing a group discussion study. *Health Policy*, 49: 63-74.
- Common, M (1997). Do existence values for cost-benefit analysis exist? *Environmental and Resource Economics*, 9: 225-238.
- Corneo, G and Fong, C(2006). *What's the Monetary Value of Distributive Justice?* CESifo Working Paper, 1706.
- Donaldson, C(1990) Willingness to pay for publicly provided goods: a possible measure of benefit? *Journal of Health Economics*, 9:103-118.
- Donaldson, C, Shackley, P and Abdalla, M (1997). Using Willingness to Pay to Value Close Substitutes: Carrier screening for cystic fibrosis revisited. *Health Economics*, 6: 145-159.
- Donaldson, C and Shackley, P (1997). Does "process utility" exist? A case study of willingness to pay for Laparoscopic cholecystectomy. *Social Science and Medicine*, 44(5):699-707.
- Dolan P, Edlin R, Tsuchiya A et al. 2008. The relative societal value of health gains to different beneficiaries, HealthTechnology Assessment Programme Journal Series Report RM03/JH11/PD, University of Birmingham
- Dolan, P., Tsuchiya, A., and Wailoo, A. NICE's citizen's council: what do we ask them, and how?. *Lancet*. 2003; 362: 918–919
- Franken, M and X Koolman(2013). Health system goals: A discrete choice experiment to obtain societal valuations . *Health Policy*, 112 (2013) 28–34
- Frey, B S., and F Oberholzer–Gee (1997). “The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out.” *American Economic Review*, 87, 746-755
- Frey, B.S. and Jegen, R. (2001) "Motivation Crowding Theory" *Journal of Economic Surveys* 15(5):589–611
- Goddard, M, Hauck, K , Preker, A and Smith, P.C (2006). Priority Setting in Health- a political economy perspective. *Health Economics, Policy and Law*, 1: 79-90.

- Hauck K, Smith PC, *et al.*: *The Economics of Priority Setting for Health Care. A Literature Review*. Washington D.C: The World Bank; 2003.
- Heyman J, Ariely D (2004). Effort for payment. A tale of two markets. *Psychol Sci*. 2004 Nov;15(11):787-93.
- Kanheman, DJ and Knetch, J (1992). Valuing public goods: the purchase of moral satisfaction. *Journal of Environmental Economics and Management*, 22: 57-70.
- Keeney RL, Raiffa H. Decisions with multiple objectives: preferences and value tradeoffs, 2nd ed. Cambridge: Cambridge University Press; 1993.
- Lindholm, L.A, Emmelin, M.A and Rosen, M.E (1997). Health Maximization Rejected. The view of Swedish politicians. *European Journal of Public Health*, 7:405-410
- Lomas, J (1997). Reluctant rationers: public input to health care priorities. *Journal of Health Service Research and Policy*, 3:103-111.
- Lenaghan, J (1999). Involving the public in rationing decisions. The experience of citizens juries. *Health Policy*, 49:45-61.
- Margolis, Howard. *Selfishness, Altruism, and Rationality: A Theory of Social Choice* . University of Chicago Press, 1982
- McDaniels, T.L (1996). “The Structured Value Referendum: Eliciting Preferences for Environmental Policy Alternatives,” *Journal of Policy Analysis and Management* 15, no. 2 (Spring 1996): 227–251
- Mitton C, Dionne F, Damji R, Campbell D, Bryan S (2011). Difficult decisions in times of constraint: Criteria based Resource Allocation in the Vancouver Coastal Health Authority. *BMC Health Serv Res* 2011, 11:169.
- Mitton, C N Smith, S Peacock, B Evoy and J Abelson (2009). Public participation in health care priority setting: A scoping review. *Health Policy*, 91(3): 219–228
- Mitton, C.R., Donaldson, C., Waldner, H., and Eagle, C (2003). The evolution of PBMA: towards a macro-level priority setting framework for health regions. *Health Care Manag Sci*. 2003; 6: 263–269
- Mooney, G (1994). *Key Issues in Health Economics*. Harvester Wheatsheaf, London.
- Murray CJL, Lauer J, Tandon A, Frenk J. Overall health system achievement for 191 countries. Geneva: World Health Organization; 2000.

Olsen JA and Donaldson, C (1998). Helicopters, hearts and hips : using willingness to pay to set priorities for public sector health care programmes. *Social Science and Medicine*, 52: 345-69

Olsen JA1 Donaldson C, Shackley P; EuroWill Group. (2005). Implicit versus explicit ranking: on inferring ordinal preferences for health care programmes based on differences in willingness-to-pay. *J Health Econ*. 2005 Sep;24(5):990-6.

Nord, E and Richardson, J (2012). Preferences for the normative basis of health care priority setting: some evidence from two count. *Health Economics* in press.

Olsen, JA, Donaldson, C , Shackley, P and the EuroWill Group (2005). Implicit versus explicit rankings: in inferring ordinal preferences for health care programmes based on differences in willingness to pay. *Journal of Health Economics*, 24: 990-996.

Sacramento Health Care Decisions, (2006). Just Coverage, Citizens Define the Limits of Health Care. November Sacramento, US.

Mataria, A, Giacaman, R, Khatib, R and Moatti, J-P (2006). Impoverishment and patients “willingness” and “ability” to pay for improving the quality of health care in Palestine: An assessment using contingent valuation method. *Health Policy*, 75: 312-328.

McDaniels, TL (1996). “The Structured Value Referendum: Eliciting Preferences for Environmental Policy Alternatives,” *Journal of Policy Analysis and Management* 15, no. 2 (Spring 1996): 227–251.

Kanheman, D, Ritov, I and Chkade, D (1999). Economic Preferences or Attitude Expressions? An Analysis of dollar responses to public issues. *Journal of Risk and Uncertainty*, 19(1-3): 203-235.

Richardson J, Wildman J, Robertson IK. A critique of the World Health Organization’s evaluation of health system performance. *Health Economics* 2003;12:355–66.

Ruta, D.A, Donaldson, C and Gilray, I (1996). Economics, public health and health care purchasing: the Tayside experience of program budgeting and marginal analysis. *Journal of Health Service research and Policy*, 1:185-193.

Schwappach, DL (2003). Does it matter who you are or what you gain? An experimental study of preferences for resource allocation. *Health Econ*. 2003 Apr;12(4):255-67.

Scuffham PA, Whitty JA, Taylor M, Saxby R. Health system choice: A pilot discrete choice experiment eliciting the preferences of British and Australian citizens. *Applied Health Economics and Health Policy* 2010;8(2):89-97.

Shafir, E, Diamond, P and Tversky, A (1997). Money Illusion. *Quarterly Journal of Economics*, 112(2): 341-74.

- Simon, H. "A Mechanism for Social Selection and Successful Altruism," *Science* .December 21, 1990. pp. 1665-1668.
- Simonsen, W and Mark D Robbins (2000), *Citizen Participation in Resource Allocation* (Boulder, CO: Westview Press, 2000
- Skedgel, C. D., Wailoo, A. J. and Akehurst, R. L. (2013), Choosing vs. allocating: discrete choice experiments and constant-sum paired comparisons for the elicitation of societal preferences. *Health Expectations*. doi: 10.1111/hex.12098
- Smith P. Developing composite indicators for assessing overall system efficiency. In: Smith P, editor. *Measuring up: improving health system performance in OECD countries*. Paris: Organization for Economic Cooperation and Development; 2002. p. 275–95.
- Stafinski T, Menon D, Marshall D, Caulfield T. (2011). Societal values in the allocation of healthcare resources: is it all about the health gain? *Patient*, 2011;4(4):207-25.
- Stafinski, T., Menon, D., and Yasui, Y (2014) Assessing the impact of deliberative processes on the views of participants: is it 'in one ear and out the other'?. *Health Expectations*. 2014 Apr;17(2):278-90.
- Stutzer, A and Frey, B (2006). Political participation and procedural utility: An empirical study. *European Journal of Political Research* 45: 391–418, 2006
- Stewart, JM, O'Shea, E and Donaldson, C et al (2002). Do ordering effects matter in willingness to pay studies in health care? *Journal of Health Economics*, 21: 585-599.
- Tversky, A and Simonson, I (1993). Context Dependent Preferences. *Management Science*, 39: 1179-1189.
- Whitty, J Peter Littlejohns (2014). Social values and health priority setting in Australia: An analysis applied to the context of health technology assessment. *Health Policy*, in press.
- Willoughby, Katherine G., and Julia E. Melkers. 1998. The state of the states: performance-based budgeting systems in 47 of 50. *Public Administration Review* 58:66–73.
- Yankelovich, D (1991). *Coming to public judgement: making democracy work in a complex world*. Syracuse University Press, Syracuse, NY.