Non-Market Resource Allocation and the Public's Interpretation of Need: An Empirical Investigation in the Context of Health Care

Jeremiah Hurley[a,b,*] Emmanouil Mentzakis[c] Mita Giacomini[b,d]
Deirdre DeJean[b,d] and Michel Grignon [a,b,e]

April 10, 2017

[a]Department of Economics, McMaster University
[b]Centre for Health Economics and Policy Analysis, McMaster University
[c]Department of Economics, University of Southampton
[d]Department of Health Research Methods, Evidence and Impact, McMaster University
[e]Department of Health, Aging and Society, McMaster University

[*]Corresponding author:
Department of Economics, KTH 132
McMaster University
1280 Main Street West
Hamilton, ON L8S 4M4
e-mail: hurley@mcmaster.ca

This research was funded by grant # 76670 from the Canadian Institutes of Health Research and by grant #2045 from the Government of Ontario through the Ministry of Health and Long-Term Care Health System Research Fund. The views expressed in this paper are those of the authors and should not be taken to represent the views of the Government of Ontario. We thank Neil Buckley, Anthony Culyer, Kate Cuff, Stuart Mestelman, Andrew Muller, and Jingjing Zhang and two anonymous referees for helpful comments, and Dave Cameron and Aleksandra Gajic for excellent research assistance. Any errors are our own. The authors have no conflicts to declare.

Abstract

The concept of need is central to the non-market allocation of many public resources, although the definition of need to serve as a basis for such resource allocation often remains contested. This study uses a discrete-choice experiment to investigate the general public's interpretation of need in the context of health care resource allocation, focusing on three commonly cited definitions of need: need as a person's baseline health status; need as a person's ability-to-benefit; and need as the amount of resources required to exhaust a person's ability-to-benefit. Analysis of participants' need judgments using a latent-class, rank-ordered conditional logit model reveals that most individuals draw on all three definitions when assessing need, and that here is heterogeneity in interpretations of need among the public. Baseline health status is the most influential and consistent determinant of need, while ability-to-benefit and resources-required-to-exhaust-benefit are considered jointly. However, while some assign greater need to those who are worse off in the sense that they have little ability-to-benefit and require large amounts of resources to achieve that benefit, others assign greater need to those who have greater ability-to-benefit and whose benefit can be achieved with small amounts of resources. The public's reasoning about need contrasts sharply in a number of ways with the types of arguments offered in the literature on needs-based resource allocation.

Keywords: resources allocation, need, discrete-choice experiment, stated-preference, latent-class model

JEL Classifications: H23, I18, C91

1 Introduction/Background

The concept of need is central to the non-market allocation of many public resources. Need, for example, underlies intergovernmental grants to achieve fiscal equalization among sub-national levels of government, and it guides both public funding to provider organizations and the provision of services to individuals in areas such as social services, education, and health care (e.g., Smith, 2007). Need also figures prominently in many theories of distributive justice, and empirical research documents that need is one of a small number of principles that consistently influences people's judgments regarding both the fair and the preferred allocation of a resource (Konow, 2003). Understanding need, therefore, is essential for assessing both the actual distribution of resources in society, especially those allocated through the public sector, and the welfare-maximizing distribution for society.

The nature of need as it relates to just social arrangements and the allocation of society's resources has been debated for centuries by philosophers, ethicists, political scientists, sociologists, economists, and others. This debate has addressed questions such as whether human needs are objective or subjective, relative or absolute, and the role of individuals versus communities in defining needs (Boulding, 1966; Braybrooke, 1987; Robertson, 1998; Hasman et al., 2006; Hope et al., 2010; Juth, 2015; Herlitz and Horan, 2016). This debate reveals widespread (though not universal) agreement that needs differ from wants and the distinction between them can be defined only in relation to community values; the presence of a need carries with it a duty by others to respond (though this duty is not absolute); need is an instrumental concept; and the relevant notion of need can differ depending on the context. Agreement, however, on the substantive characteristics that determine need in many concrete allocation contexts remains elusive.

This elusiveness and the non-welfarist foundations of most conceptions of need generates unease within modern, welfarist economics, especially claims of need for specific goods, which many economists view merely as attempts to “dress-up” wants and preferences to give them greater normative force. The notion of “basic human needs”, however, underlies aspects of welfare analysis such as the definition of the poverty line and social indicators in areas such as housing, food, and development (e.g., the UN Development Index). This work typically begins from the view that all humans require a minimum level of a small number of basic material goods—food, housing, etc. —to achieve a basic level of well-being, and that when access to these goods falls below the minimum standard, a person has a need for one or more of these goods. The amount of need can be gauged by the deviation between an individual’s consumption of these goods and the minimum standard.

While empirical studies document the importance of need for individuals’ judgments regarding their preferred and the just, or fair, allocation, these studies provide limited evidence regarding how the public judges need itself. Such studies typically present participants with a scenario that involves an allocation dilemma, and then asks respondents a question that elicits the desired information. Some studies (e.g., Hurley et al., 2011) ask participants to rate the importance of different principles (e.g., need, efficiency, merit), without defining the principles, which confirms the importance of need as an allocation principle but provides no insight as to how the participants judge need. Other studies (e.g., Skitka and Tetlock, 1992) present participants with a specific definition of need (e.g., severity of illness in a health care allocation dilemma) and elicit judgments of its importance for allocation. While confirming the importance of the identified factor, we do not know whether the presented definition corresponds to the participant’s own definition of need. Still other studies (e.g., Konow, 2001; Cappelen et al., 2013) present information on an attribute the researcher believes represents need, without making any explicit reference to need, and then infer the importance of need based in the influence of that attribute.

Such studies often use broad indicators of basic needs (living in a developing vs. developed country; living at subsistence level vs. above subsistence; income below poverty line vs. above poverty line) that are not likely to be highly disputed but which provide limited guidance to good-specific allocation problems in sectors such as housing or health care for which there may be less consensus regarding need for specific goods.

This study seeks to fill this gap in the literature by investigating directly, in a specific context—the allocation of health care—how the public interprets need. After presenting participants a scenario pertaining to the allocation of pain-relief pills among individuals experiencing pain, rather than ask participants either their preferred or the fair allocation of pills, we ask them which individual has the greatest (least) need for the pills. We investigate this question in a health care context of for a number of reasons. First, the public strongly supports allocation of health care according to need as both the fairest way to divide limited health resources (Hurley et al., 2011; Yaari and Bar-Hillel, 1984) and the preferred basis for allocating health care resources (e.g., Abelson et al., 1995; Shah, 2009; van Exel et al., 2015); allocation according to need is the stated objective of public health systems around the world (van Doorslaer et al., 1993); and need figures prominently in many ethicists' analyses of the just basis for allocating health care (e.g., Daniels, 1985, 2008) and in many health economists' assessments of efficiency and equity in the health sector (Williams, 1978; Evans, 1984; Culyer, 1995). Second, health care is fiscally important: in many countries health care constitutes the largest single public-sector program and approximately ten percent of all economic activity. Third, questions of need in health care arise at a micro level for specific treatments and conditions. Agencies such as the English National Institute for Health Care Excellence (NICE) advise health care authorities regarding funding for technologies and treatments based in part on public input regarding priorities, and there is a large and growing literature on health care priority-setting whose goal is to inform funding and resource allocation within health systems. In no other sector is the link between such research and decision-making as extensive, deep and direct. This study draws on this extensive literature on health care priority-setting (for recent reviews see, e.g., Whitty et al., 2014; Gu et al., 2015), but differs from it by directly asking participants about their assessment of need rather than about their preferred allocation. Finally, within health care, debate about the relevant concept of need to underpin resource allocation has focused on three specific definitions of need (with many minor variants)—need as baseline health status, need as ability-to-benefit, and need as the amount of resources required to exhaust benefit—that lead to quite different resource allocations (Culyer and Wagstaff, 1993). In the absence of a consensus or an understanding of how the public interprets need in the context of resource allocation, the practical implications of the public's support for allocation according to need is unclear.

These competing definitions of health care need incorporate elements not found in broader analyses of basic needs, a difference rooted in distinctive characteristics of health and health care. A necessary condition for a person to have a need is that they experience a deficit relative to a norm, e.g., relative to the basic minimum of food, housing or income. The definition of need as baseline health status reflects this logic, where full health serves as the reference standard against which to judge the deficit. Within this framework, the magnitude of need is proportional to the size of the deficit. This makes sense for food, housing, or income, for which the effectiveness of the resources used to address the need (housing, income) is both clear and common across individuals.[1] But the use of baseline health status as a measure of need for health care is contested in part because this

[1]Nuances, of course, do arise. The resources required to address a housing need for a physically disabled individual differ from an able-bodied individual. But because these claims pertain to shared, basic material human needs relative to a common material standard within a society the material resources required and their effectiveness are, except for unusual circumstances, similar across individuals. Further, such measures abstract from preferences—a homeless person may prefer to not live in a traditional dwelling—but these preferences do not alter the assessed need for housing.

relationship between the size of the deficit and the resources required to overcome the deficit breaks down. Some health conditions have no effective health care treatment: no amount of health care will reduce the deficit. For consequentialists, a necessary condition for a health care need to exist is that there be an effective treatment; the mere existence of a health deficit is not a sufficient basis for a claim of need (for health care at least) (Williams, 1978; Culyer, 1995). Such reasoning leads to the second definition: ability-to-benefit. The amount of need is directly related to one's ability-to-benefit from health care treatment. This definition aligns more closely within economic reasoning, and specifically extra-welfarist approaches in health economics (Hurley, 2000). But it suffers from an important limitation as a guide to resource allocation: the magnitude of the ability-to-benefit is often a poor indicator of the resources required to achieve the benefit. An otherwise healthy person in anaphylactic shock from an allergic reaction has the greatest possible ability-to-benefit—immediate death vs. full health—yet the only resource required is an inexpensive epi-pen. In contrast, a person with a hernia who can live a largely normal life with but a few restrictions may require expensive surgery to return to full health. This limitation leads to the third definition: the amount of resources required to exhaust benefit using the most cost-effective treatment. This definition corresponds more directly to the problem faced by a funder: determining the resources required to address the extant burden in the population. Setting the reference standard to define the amount of need as the point at which the marginal product of health care is zero may strike economists has puzzling, but it again reflects special issues that arise in the health context. In sectors beyond health care, such as housing, need can be defined in terms of the resources required to bring an individual up to the common, basic standard, which is in principle achievable for everyone (given enough resources). But in health care, it is often not possible to bring everyone back to the reference standard of full health, no matter how much health care is provided. Hence, full health cannot serve as the reference standard for assessing resource need. But health care treatment is subject to diminishing returns, so the point at which the marginal product of treatment becomes zero can serve as a common reference point to measure need across illnesses and treatments. These complexities associated with assessing need in health care make it especially interesting to investigate hows the public's interpretation of need corresponds to these definitions developed by health ethicists and health economists.

We use a discrete-choice experimental design to investigate support among the general public for each of these three definitions of health care need. Participants are presented with hypothetical scenarios involving three individuals characterized according to their health status, ability-to-benefit from health care, and the amount of health care required to exhaust benefit. Respondents are then asked which person in the scenario has the *greatest need* and which has the *least need.* Drawing on multi-criteria approaches to justice and fair resource allocation (e.g., Miller, 1976; Deutsch, 1985; Frohlich and Oppenheimer, 1992; Scott et al., 2001; Konow, 2003), which argue that people use multiple foundational justice principles (weighted differently across contexts) when making fairness judgments, we hypothesized that people's judgments of need would draw on all three definitions. This idea also has direct parallels to multi-criteria decision analysis advocated for health technology assessment (e.g., Thokala, 2012), and to conditional allocation rules and trade-offs among principles investigated in Schwettmann (2009) and Ahlert (2013). A discrete-choice design is well-suited for testing the hypothesis that individuals draw on multiple criteria when making judgments of need, and enables us to estimate the weights placed on each while admitting as a special case the possibility that a person always bases judgments on a single need criterion.

The study contributes in two ways to the growing empirical social choice literature (e.g., Gaertner and Schokkaert, 2012) that investigates public support for ethical concepts and principles heretofore subject solely to conceptual analysis and argument. Our results show that the public thinks about need quite differently than has much of the academic analysis of need, highlighting the importance in debates about resource allocation to get beyond high-level labels such as "need" to the substantive underlying interpretations. Second, for the choice analysis we extend the latent-class family of econometric estimators for choice models by implementing a rank-ordered conditional logit specification incorporating unobserved heterogeneity and panel data structures. To our knowledge, this is a first application of this estimator, which has potentially widespread application in the analysis of social choice based on individual rankings over alternatives.

2 Methods/Experimental Design

We used a stated-preference, discrete-choice methodology in which the choice scenarios include individuals described by attributes associated with each of three above-noted prominent definitions of need: baseline health status, ability-to-benefit from health care, and the amount of health care resources required to exhaust benefit. The discrete-choice experiment randomized the combinations of these attributes across the choice scenarios, thereby allowing us to estimate the weight subjects place on each definition when deciding who among a group of individuals has the greatest need.

2.1 The Survey: Choice Scenarios and Attributes

The survey questionnaire, which was pilot tested and revised as appropriate to ensure clarity, included three parts: an introduction and scenario description; a set of choice problems; and a short demographic survey. The introduction noted that the health care systems of many countries around the world, including Canada, are designed to provide services to individuals based on their need for care; that it can be difficult for providers and health systems managers to assess people's need for care; that there is no universal agreement on what is most important in determining an individual's need for care; and that the purpose of the survey was to gain an understanding of the respondent's views regarding people's need for care.

The choice scenario considered three individuals who suffer from chronic pain. The intensity of pain experienced by the three individuals is identical and is sufficient to prevent them from participating in many normal daily activities. The individuals differ, however, in the amount of time that they are in pain each day without treatment and in the hours of pain relief they can obtain from pain medication. Even with medication, some individuals are not able to obtain complete relief from pain. The pain medication, which has no negative side effects and can safely be taken in the amounts considered, is available to the individuals at no cost. The supply of pain medication, however, is insufficient to provide pain relief to all individuals. The individuals are described as identical in all respects not explicitly noted, such as age, sex, income, marital status.

The basic design of the choice scenario follows closely that used in studies of equity in the allocation of resources (Yaari and Bar-Hillel, 1984; Konow, 2003). We frame the scenario in terms of chronic pain and pain-relief medication for a number of reasons. Both baseline health status and the health care resource had to be easily quantifiable, and it had to be plausible that the amount required would vary across individuals. Using chronic pain as the underlying health condition allows us to quantify baseline health status as the hours of pain each day with no medication, and chronic pain is commonly treated with medication, which is readily quantifiable as the number of pills required and plausibly varies across individuals. Finally, pain does not carry stigma, both pain and pain medication are familiar to most individuals, and pain/pain medication has been used previously in studies of the role of need in health care resource allocation (Kahneman and Varey, 1991; Cuadras-Morato et al., 2001; Hurley et al., 2011).

The choice scenarios presented three attributes regarding the individuals' pain and ability to obtain pain relief from the medication available.

1. Number of hours each day free of pain with no treatment. This attribute represents baseline health status and describes the number of pain-free hours per 24-hour day that an individual would experience if they took no pain medication. It had four values: 0, 4, 8 or 12 hours. A value of 4 hours, for example, means that if an individual took no pain medication they would be in pain 20 hours each day and free of pain 4 hours.

2. Number of hours of pain relief possible from medication. This attribute represents ability-to-benefit and describes the maximum number of hours of pain relief an individual could obtain by taking pain-relief pills. This pain relief is in addition to any pain-free hours they experience with no medication. The attribute had four values: 4, 12, 16 or 24 hours. A value of 12 hours, for example, means that the maximum number of hours of additional pain relief possible by taking medication is 12 hours per day.

3. Number of pain-relief pills required to obtain the maximum hours of pain relief from medication. This attribute represents the amount of resources required to exhaust benefit and describes the number of pain-relief pills an individual would need to take each day to obtain their maximum possible number of hours of pain relief per day. The scenario stated that, because of biological differences among the individuals, the effectiveness of the pills differs across the individuals. It had four values: 2, 6, 8 or 12 pills. A value of 6 pills, for example, means that the individual must take 6 pills per day to obtain the maximum number of hours of pain relief possible.

2.2 Experimental Design

A full-factorial generic design with three four-level attributes generates 64 possible combinations. Allowing for two-way interactions between baseline hours of pain and hours of pain relief possible and between baseline hours of pain and pills required to exhaust benefit, a fractional-factorial design was produced with 16 choices scenarios of three alternatives (Kuhfeld, 2005; Zwerina et al., 1996). To enable us to test for response consistency, two extra choice sets were added and each was repeated within the survey (separated by at least nine intervening questions). Participants therefore responded to a total of 20 choice scenarios with three alternatives each. The design took account of the dependence among the attributes. For an individual, the maximum possible additional hours of pain relief per day (attribute 2) is 24 less the number of hours free of pain each day without medication (attribute 1). Hence, only 48 of the 64 combinations of the full factorial are feasible (e.g., baseline of 12 pain-free hours each day and a maximum hours of pain relief equal to 16 is not feasible), which required that we place restrictions on the experimental design. Further, the value of attribute 3 (number of pills required to exhaust benefit) equals attribute 2 divided by the hours of pain relief obtained per pill. To ensure independent variation between attributes 2 and 3, the pain relief obtained per pill varied across individuals. All aspects of the experimental design were performed using SAS 9.1.3 built-in capabilities (Kuhfeld, 1997). Figure 1 presents an example of the scenarios presented to participants. The on-line Appendix 2 presents an abbreviated version of the full survey.

2.3 Survey Administration

The survey was administered to a sample drawn from the community-dwelling population of four english-speaking provinces of Canada—British Columbia, Saskatchewan, Ontario, and Nova Scotia—that represent key regions of the country. It was administered using a mixed-mode methodology in which community participants were recruited using a letter of invitation sent via regular mail, which allowed us to use postal contact information (obtained from a marketing research firm) for a representative, random sample of the provincial populations, but participants completed the survey via the Internet (Gajic et al., 2012). Participants were compensated for their participation via a lottery for cash prizes of $250 (Cdn). The study was approved by the McMaster University Research Ethics Board.

3 Econometric Analysis

For each participant, for each of the 16 choice scenarios involving three individuals in need, we have data on the individual the participant rated as having the *greatest* need and the individual the participant rated as having the *least* need, which allows us to determine a participant's full ranking of the individuals with respect to level of need. To exploit this information, we developed a latent-class rank-order logit model, which accounts for unobserved individual preference heterogeneity. Such heterogeneity implies that individual preferences cannot be all depicted by a single underlying parameter vector and that behavioral patterns not explained by observed characteristics must be captured in via the estimation approach. Latent-class models are increasingly employed in the discrete-choice experimental literature (Greene and Hensher, 2003; Mentzakis et al., 2011) because, in addition to accounting for such unobserved individual heterogeneity, they allow for partial relaxation of the assumption of independence of irrelevant alternatives (IIA) and accommodate the panel structure of data. In the latent-class model, individuals are implicitly sorted into a set of C classes, where each class represents a different data-generating process (i.e., in our setting, a distinct pattern of weighing the three need-related attributes when judging need). It is termed a latent-class model because the class of any specific individual is determined probabilistically and hence remains unknown to the analyst (Greene and Hensher, 2003). Assuming an additive deterministic component ($V_{iq} = \sum_{k=1}^{K} \beta_k X_{iqk}$) and a stochastic component (ϵ_{iq}), estimation is based on a random utility model where individual utility is $U_{iq} = V_{iq} + \epsilon_{iq}$. Further, assuming that ϵ_{iq} is independent and extreme value type I (EV1) distributed, the rank-order specification can be viewed as a series of sequential conditional logits (McFadden, 1974), where at each point in the series the individual chooses the most preferred alternative out of the remaining ones (Luce, 1959).[2] Hence, the probability that individual q will rank alternative i the highest in choice scenario t conditional on falling within class c is

$$P_{iqt|c} = Pr(i > max(1, ..., J)) = \frac{e^{X_{iqt}\beta_c}}{\sum_{j=1}^{J} e^{X_{jqt}\beta_c}} \tag{1}$$

The remaining alternatives are also ranked following a conditional-logit probability, where the probability that alternative m is ranked highest of the remaining alternatives is

$$P_{mqt|c} = Pr(m > max(2, ..., J)) = \frac{e^{X_{mqt}\beta_c}}{\sum_{j=1}^{J} e^{X_{jqt}\beta_c} - e^{X_{iqt}\beta_c}} \tag{2}$$

Eqs. 1 and 2 embed an implicit assumption that the weights placed on the attributes when ranking individuals do not change as we move down (or up) the rank. As noted above, class assignment in the latent-class model is probabilistic, so following Greene and Hensher

[2]In our setting, participants are first asked who has the greatest need, so the choice set for this decision is all three alternatives; they are then asked who has the least need, so the choice set for this decision is the two remaining alternatives.

(2003) let H_{qc} be the probability that an individual q falls in class c:

$$H_{qc} = \frac{e^{Z_q \gamma_c}}{\sum_{c=1}^{C} e^{Z_q \gamma_c}} \tag{3}$$

where Z_q is a set of individual traits/covariates that influence the probabilities of membership in the different classes. Class distinctions are empirically determined by the underlying data and, as such, there is no need to impose *a-priori* interpretations on the classes or ascribe behavioral patterns to specific groups. Parameterization of class probabilities with individual traits is undertaken only if it improves model fit; it is possible that observed covariates do not characterize the latent patterns embedded in the data. From the above, the contribution of individual q to the likelihood is

$$L_q = \sum_{c=1}^{C} H_{qc} \prod_{t=1}^{T} P_{iqt|c} \times P_{mqt|c} \tag{4}$$

and the latent-class rank-order log-likelihood for sample size N becomes

$$\ln L = \sum_{q=1}^{N} \ln \left[\sum_{c=1}^{C} H_{qc} \prod_{t=1}^{T} P_{iqt|c} \times P_{mqt|c} \right] \tag{5}$$

Our experimental design did not allow for ties or incomplete rankings, so no such considerations were incorporated into the estimation (Allison and Christakis, 1994). The number of latent classes is determined exogenously based on the model fit to the data as measured by criteria such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) (Greene and Hensher, 2003; Swait and Adamowicz, 2001). The latent-class rank-order logit estimation routines were written in Stata.[3] In order to verify that the proposed estimator is identified and performs well, we tested its properties using simulation

[3]Estimation is performed using both the ML and EM algorithms.

experiments, which confirmed its good performance (see Appendix 1 for details).

4 Results

4.1 Sample Descriptive Statistics

The sample includes 349 community participants (Table 1), which reflects a response rate of 12% for completed surveys, which is lower than traditional mail surveys but consistent with response rates common for general population web surveys (Dillman, 2009; Gajic et al., 2012; Manfreda et al., 2008). The sample has a mean age of 53.5, 60 percent are males, 55 percent report excellent or very good health, 73 percent are married, 78 percent are post-secondary graduates (university, college, professional/trade-school), 57 percent have full-time jobs, 89 percent own their house, and household income ranges from less than $20,000 to over $100,000, with the median falling between $50,000 and $100,000.

For comparison, we include characteristics of the community dwelling populations (derived from the Canadian Community Health Survey [CCHS]) of the provinces included in our survey. Our sample is somewhat older, more male, more likely to be married, more likely to own their home, and more likely to be working; it is comparable to the population with respect to educational attainment and health status (0.55 vs. 0.59 in excellent or very good health). Our sample has a higher proportion who do not provide information on their household income, and income appears to be missing disproportionately from both high and low income households. The proportion with household income of $50k-$99k matches the CCHS and is consistent with the Census a median family income of $77k). The higher proportion of male respondents and slightly higher proportion of post-secondary graduates is consistent with the evidence (Dillman, 2009) that educated males tend to respond more to on-line surveys than do females or less educated individuals. The slightly older age of respondents may reflect the fact that older (retired) individuals likely have a lower average value of time and may have higher-than-average interest in health care issues. While our sample is not strictly representative of the community-dwelling population, it contains a cross-section of the population and is broadly comparable on a number of characteristics. Given that our purpose is not to obtain precise estimates of population parameters, but to investigate patterns of judgments among the public regarding need and how these patterns correspond to the conceptual definitions of need, the sample contains a sufficient cross-section of the public to provide insight into such patterns.

4.2 Analysis of Judgments of Need

4.2.1 Model Choice

The latent-class model revealed important preference heterogeneity and, compared to the standard rank-ordered logit, provided considerable improvement in model fit. We compared latent-class models with two to five classes based on BIC goodness-of-fit statistics. The four-class specification fit better than the two- and three-class models and each class represents a distinct pattern of need judgments. Allowing for a fifth latent class neither improved model fit meaningfully nor generated a qualitatively distinct pattern of need judgments beyond those in the four-classes model. We therefore present the results based on rank-ordered latent-class logit specifications with four latent classes. We compared 4-class models with no interaction terms and with two-way interaction terms between baseline health and each of ability-to-benefit and resources-required-to-exhaust benefit. The two-way interaction terms were generally not significant and the key results are very similar across the models with and without interaction terms. Finally, we compared 4-class models estimated over the full sample and over the sub-sample with consistent responses. To test for response consistency we define consistent responders as those who chose the same alternative as representing greatest need for both of the repeated choice scenarios. By this definition, the preferences of 212 participants are consistent.[4] A logit analysis of the probability of being consistent as function of individual demographic characteristics revealed no significant associations with the exception of education—those with less education were more likely to report inconsistent responses than those with high levels of education. While there are some differences in the magnitude of coefficients and class sizes for the models estimated on the two samples, there is considerable correspondence between the two sets of results: the same class types with parallel interpretations emerge for both samples. However, as one would expect, the size of class that corresponds to those with the weakest preferences over the definitions and that represents a type of residual category of individuals without strong preferences, is larger in the full sample. Below we present results for the model with no interaction terms estimated over the consistent sample.[5]

4.2.2 Regression Results

Table 2 presents results for the four-class specification. The estimated class shares are 0.378, 0.189, 0.337, and 0.096, so each constitutes a meaningful proportion of the sample. We had no *a priori* expectations that judgments of need would be systematically associated with demographic and socio-economic characteristics, and a test for associations between measured individual characteristics (age, sex, education, work status, income) and class assignment revealed few associations. The coefficients associated with each class are interpreted identically to those for a standard conditional logit model. For example, the

[4]Of those classified as inconsistent, only 10% missed both; 90% missed only a single test question.

[5]Results for the full sample and for models with interaction terms are listed in on-line Appendices 3 and 4 for the review process.

coefficient of -2.968 on “Pain Free 4 hours” in class 1 indicates that, other things equal, participants judge a person pain-free for 4 hours per day without medication as being in lesser need than a person in the reference category of pain-free for 0 hours per day without medication. The reference category for each attribute is the attribute value that represents the greatest need according to the associated definition, so negative coefficients indicate judgments of need consistent with that definition; positive coefficients indicate an unexpected pattern in which a person with the indicated level was judged to be in greater need than a person in the reference category.

A few patterns are notable. All three attributes influence judgments of need for all classes except class 4, for which only ability-to-benefit is significant. Baseline health influences judgments of need in the expected manner for all four-classes, though the effect is small and not significant for class 4. But each of ability-to-benefit (ATB) and resources-required-to-exhaust-benefit (RREB) influence judgments of need in an unexpected manner in two classes: for classes 1 and 3, need is judged to be a decreasing function of ATB, and for classes 2 and 4, need is judged to be a decreasing function of RREB (though the relationship is not statistically significant for class 4).

To provide a better sense of the role of each of these attributes in judgments of need, based on the results presented in Table 2, Figures 2-5 plot the probability that a given combination of attribute levels is judged to represent greater need than the reference combination of attributes: pain-free 0 hours per day without medication, maximum possible pain relief of 24 hours, and 12 pills required to obtain the maximum pain relief.[6] This reference, or baseline, combination includes the greatest level of need posited by each of the three definitions considered, so if participants subscribe in the expected manner to a weighted combination of the three definitions, this baseline combination should have a

[6] All probabilities are conditional probabilities, i.e., conditional on falling into the noted class.

higher probability of being judged the greatest need than any other combination of attribute levels. In each figure, the vertical axis plots probability; the horizontal axis plots each of the 48 feasible combinations of the attribute levels. Beginning at the origin, the first 16 combinations on the horizontal axis are those in which a person is pain-free 0 hours per day; the next 12 combinations are all those in which a person is pain-free 4 hours per day; the next 12 are combinations in which a person is pain-free 8 hours per day; and the final 8 are combinations in which a person is pain-free 12 hours per day. Possible pain relief and number of pills vary within each of these groupings. Within the 16 combinations for which a person is pain-free 0 hours per day, the first 4 represent an ability-to-benefit from medication of 4 hours; the next 4 represent an ability-to-benefit of 12 hours, and so forth. Finally, the number of pills required to exhaust benefit varies within each subset of 4 combinations that represent a given level of baseline pain and ability-to-benefit from medication. As noted, the reference combination for all probability calculations is pain-free (PF)=0, ability-to-benefit (ATB)=24, pills required to exhaust benefit (RREB) =12.

Figure 2 presents the results for class 1. The steep negative left-to-right gradient across the figure reflects the dominant role of baseline health status in judgments of need for this class. Individuals in pain 24 hours a day without medication (first 16 observations) have probabilities of being judged most in need as high as 0.92 and never less than 0.20; in contrast, individuals with 8 or 12 hours pain free without medication have almost zero chance of being judged in greater need (regardless of their values for ATB and RREB) than individuals with the reference combination of attribute values. ATB and RREB exert a meaningful influence on judgments of need only among those with low-baseline health. The effect of ATB is revealed by examining the four clusters of four observations within each level of baseline pain. Focusing first on individuals pain-free 0 hours per day (obs 1-16), we see a negative gradient moving from the cluster with 4 hours ATB (obs 1-4) to the cluster with 24 hours of maximum benefit (obs 13-16). This is unexpected: need is decreasing rather than increasing in ATB. The effect of RREB is revealed by the pattern for the four observations within each cluster of baseline health and ability-to-benefit. Focusing on the first four combinations (obs1-4), we see the expected positive relationship between need and RREB. For class 1, the combination of attributes with the highest probability of being judged most in need is obs 4: 0 hours pain-free, maximum of 4 hours of benefit possible, and 12 pills required to achieve that benefit. Such an individual is, in a sense, in the worst possible situation: they experience the most pain possible, benefit little from medication, and require a lot of medication to achieve this small benefit.

Figure 3 presents the results for Class 2. Once again, the sharp left-to-right negative gradient as baseline health increases confirms baseline health as the dominant factor influencing judgments of need. In contrast to Class 1, however, we now have the expected positive relationship between need and ATB (i.e., the upward trend among the clusters for each level of baseline health), but an unexpected negative relationship between need and RREB (i.e., the negative gradient across observations within each 4-observation cluster given baseline health and ATB). Given this pattern of judgments, the combination of attributes with the highest probability of being judged most in need is obs 13: 0 hours pain-free, ATB of 24 hours, and only 2 pills required to achieve complete pain relief. Such an individual has low baseline health without medication but benefits greatly from medication and requires only a small amount of medication to achieve this large benefit.

Figure 4 presents the results for Class 3. The absence of a sharp left-to-right gradient indicates that all three factors influence judgments, with no single factor dominant. To see the large impact of baseline health, compare the first, leftmost observation (0, 4, 2) when pain-free hours equals 0 to its corresponding observations when pain-free hours equal 12 (12, 4, 2) — the probability of being judged in greater need than the reference combination has fallen from 0.90 to approximately 0.15, reflecting the large impact of baseline health. But unlike classes 1 and 2, ATB and RREB exert important influence even at high levels of baseline health—at least one combination of attributes with 12 pain-free hours (12, 4, 12) has a probability of over 0.50 of being judged in greater need than the reference combination. As for class 1, we again see an unexpected negative relationship between need and ATB, and an expected positive relationship between need and RREB. The distinguishing feature in judgments between Class 3 and Class 1 is the more equal weight given to each attribute in Class 3 compared to the dominant influence of baseline health in Class 1.

Finally, Figure 5 presents the results for Class 4. The most striking feature is the relatively small influence of each of the attributes. All of the probabilities fall within a relatively narrow band between 0.15 and 0.55. The shallow gradient as one moves left to right across the diagram is consistent with the small, non-statistically significant coefficients on baseline health in Table 2. Similarly, there is no meaningful gradient within each 4-observation cluster, consistent with the small, non-statistically significant coefficients on RREB. The most prominent pattern is the positive, modest gradient between need and ATB.[7]

In summary, most participants draw on all three concepts when assessing need. Baseline health is the only attribute that consistently affects judgments in the expected direction, and is the overwhelmingly dominant factor in two of the four classes. The influence of ATB and RREB is more complex: each influences judgments in the expected manner for two classes and in an unexpected manner for two classes. In combination, they generate two distinct patterns: for Classes 1 and 3, conditional on baseline health, need is a decreasing function of ATB and an increasing function of RREB, so that the individuals judged most

[7]Results for class 4 in the full sample differ slightly: the gradient for baseline health is similarly shallow and the effects for ATB and RREB modest, but there is a modest negative relationship between need and ATB and a positive relationship between need and RREB. Such differences associated with the inclusion of inconsistent individuals suggests that the shallow gradient for baseline health is the key distinguishing characteristic of Class 4.

in need are those with poor baseline health, little ability-to-benefit from treatment, and who require a large amount of resources to obtain this small benefit. Need is associated with being "worst off, " a situation that is highly intractable. In contrast, for Classes 2 and 4, conditional on baseline health, need is an increasing function of ATB and a decreasing function of RREB, so that the individuals judged most in need are those with poor baseline health, large ability-to-benefit, and who require small amounts of resources. Conditional on baseline health, need in this case is associated with tractability, effectiveness and efficiency — those for whom the resources will generate the most benefit are most in need.

We had hypothesized this would be the case and chose a DCE design to be able to identify such a phenomenon. But the ways in which people draw upon and combine information on ATB and RREB is unexpected. This pattern is not the result of the mathematical dependence among these two attributes: the experimental design allowed the levels of these two attributes to vary independently.[8] Although these patterns potentially reflect well thought out, reasoned responses by participants, they raise the question of whether participants misunderstood some aspect of the survey. To investigate this further, we conducted an additional sub-study that included a qualitative component to answer these questions and understand better the reasoning of participants.

5 Qualitative Follow-up Study

The subsequent sub-study collected two additional types of qualitative information: written feedback regarding why participants made the choices they did and data from a semi-structured interview with a random sample of participants in this follow-up study.

[8]Tests for independence (based on Spearman's rank correlation and Kendall's Tau statistic) between the levels of these attributes specified in the survey scenarios failed to reject the null hypothesis of independence (p = 0.41 and 0.43 respectively).

5.1 Design of the Qualitative Study

All participants in the sub-study provided responses to the same set of choice scenarios as in the full study. However, to gain insight into the reasoning of participants, for a subset of five questions, after a participant chose the individuals they judged to be most and least in need, they were asked the following open-ended question, "Please explain why you believe this individual has the [least/greatest] need." Because inclusion of this open-ended question could affect their judgments, the follow-up study included two arms to enable us to test for any such effects: one arm received the original survey and the second arm received the survey with the open-ended questions. In addition, to gain further insight into participants' judgments, we conducted debriefing interviews with a random subset of participants. The debriefing interview probed participants on five topics: their understanding of the DCE survey/exercise, including any points of confusion or difficulties they had in completing the survey; what need, and health care need specifically, meant to the participant; for each of two scenarios, why they judged the chosen individual as having the greatest/least need; how important each of the three attributes (BH, ATB, RREB) was to them in general for assessing need; and what else they would like to have known to assess need (see on-line Appendix 5 for the interview instrument).

Because the sub-study included in-person debriefing interviews, the sample was drawn locally from students and staff of McMaster University, and the survey was administered electronically in the McMaster Experimental Economics Laboratory (McEEL). A total of 106 individuals participated in the follow-up study, 53 completed the original survey and 53 completed the modified survey that included open-ended questions regarding their choices. A subset of 22 randomly chosen participants participated in debriefing interviews.

5.2 Results

5.2.1 Quantitative Analysis

Analysis of responses to the follow-up survey revealed that: (a) responses did not differ between those who completed the original survey and those who completed the modified survey with the open-ended questions; and (b) responses did not differ in a meaningful way between the follow-up university-based sample and the original community-based sample (see on-line Appendix 6 for results for the follow-up sample).

5.2.2 Qualitative Analysis

The debriefing interviews confirmed that participants understood the nature of the exercise and the questions they were asked in the survey, and participants expressed no difficulty in completing the survey. Participants were articulate in both the written responses and interviews regarding their judgments, approached the judgments of need systematically and thoughtfully, and could readily explain their judgments when asked to do so in the debriefing interviews. Both the written comments and interview responses confirm that the unexpected pattern of judgments do not reflect confusion or lack of understanding by participants but rather are deliberate and based on reasoned responses.

The written responses and the debriefing interviews confirmed that, for a large proportion of individuals, baseline health was the primary consideration when judging need. A large majority of participants looked first at baseline health, and then considered ATB and RREB only when baseline health was not decisive.

> "I always go with the hours each day free with pain with no treatment."(#21)

> "So automatically if they had less hours pain free I put them ahead of, like, B who had 12 hours." (#13)

Most participants considered ATB and RREB jointly, usually creating a synthetic variable expressed either as hours of pain relief per pill (ATB/RREB) or a measure of dosage (RREB/ATB).

> "I would look at the amount of pills that are required to obtain the maximum amount of hours of pain relief and then I would figure out a ratio. Like how many pills give each patient the most amount of hours." (#6)

> "I looked at the last two kind of together, I didn't isolate the two"(#4)

Having generated these synthetic variables, however, participants differed on how the variables influenced judgments of need. Consistent with the quantitative findings in both the community and follow-up samples, one group interpreted more pain relief per pill as a signal of greater need:

> "So I mean for the least amount of resources he would get just as much pain relief." (#2)

But another group interpreted less pain relief per pill as as a signal of greater need:

> "So individual C needs them more because [yeah] they would need more pills [yeah] for the same amount of pain relief." (#10)

It appears that those who emphasized tractability and efficiency tended to focus more on the resource constraint described in the introduction to the survey. A number of interviewees who appealed to this efficiency criterion when making judgments mentioned the issue of resource scarcity, while none of the interviewees who appealed to the converse notion of intractability mentioned the issue of resource scarcity.

In summary, the unexpected pattern of findings with respect to ATB and RREB reflect an accurate understanding of the choice problem, and consistent, reasoned responses by participants. When important to their judgment, most participants considered ATB and RREB jointly. This joint consideration, however, led to contrasting judgments of need between those who focused on tractability and appealed to notions of efficiency and those who focused on intractability and responding to those in a bad situation.

Given the qualitative evidence that many participants explicitly used the ATB to RREB when making judgments, using the community sample we estimated a restricted model that included only baseline health and the ratio of ATB to RREB (rather than entering each separately) (Table 3). We specified the ratio categorically as less than 1 (reference category), 1 to 2, 2 to 4, and greater than 4. The fit is almost identical to the unrestricted model: the 4-classes model again fits best, the shares allocated to each class are nearly identical to the unrestricted specification, the coefficients on baseline health are very similar, and we see the expected patterns across the four classes: for two classes need is negatively associated with the ratio and for two classes it is positively associated with the ratio.

6 Discussion

Although the concept of need has been investigated in a number of allocation contexts, this study offers new insight into the public's interpretation of need in the context of resource allocation. As hypothesized, in the health care context, most individuals draw to some extent on all three definitions when making judgments of need using a multi-criteria approach in which each factors exerts an influence. Such a multi-criteria approach to principles-based assessments and decision-making has important implications both for resource allocation itself and for research on resource allocation. It suggests, for instance, that commonly used methods in empirical research on resource allocation that ask a participant to choose one among competing allocation principles may force a choice quite different from an individual's preferred approach, which would combine principles in a weighted fashion. It also highlights the importance for resource allocation of understanding how the public perceives the relationship among principles.

Our findings document substantial heterogeneity in individuals' interpretations of need. The latent-class analysis identified four distinct patterns in judgments of need distinguished by the weight placed on each of the three definitions of need investigated and by the nature of the relationship between the determinants and need. Baseline health status is the most consistent and influential determinant of people's judgments: it has the broadest support and, for the majority of individuals, is quantitatively the most important factor affecting their judgments. Baseline health status is the only definition that consistently influenced judgments in the expected manner (i.e., better baseline health status is associated with less need). Many participants considered ability-to-benefit or resources-required-to-exhaust benefit only if baseline health was, in a sense, not decisive. These findings are perhaps not surprising as baseline health is the definition that most closely corresponds to the central idea of need as the size of a deficit relative to a standard, which dominates need definitions in other sectors (food, housing, income). The public's emphasis on a person's baseline health status, however, contrasts with definitions of need advocated by health economists who argue that, while a health deficit is a necessary condition for the presence of a need, because some health conditions have no effective treatments and the size of the health deficit can be a poor indicator of the resources required to achieve benefit, definitions of health care need must emphasize these additional considerations (and even ignore size of the baseline health deficit). When participants considered ability-to-benefit and resources-required-to-exhaust-benefit, they tended to do so only jointly through the consideration of their ratio.

But the way this information influenced judgments differed sharply across individuals. For one group of individuals, conditional on baseline health status, lesser ability-to-benefit and/or large resource requirements signaled greater need and an imperative to respond to those in a difficult, intractable situation. For the second group, whose reasoning emphasized notions of effectiveness and efficiency, those with greater ability-to-benefit and/or requiring fewer resources had greater need, implying that, conditional on baseline health, allocation according to need would generate the most benefit with the resources. Each of these approaches reflects its own internal logic, but they have radically different implications for allocations that would result from applying the principle of allocation according to need, and particularly the relationship between the principle of allocation according to need and efficient allocations. They also highlight a degree of conflict that can arise even within multi-criterion approaches designed to reconcile differing approaches to allocation judgments. While multi-criterion approaches commonly find differing weights places on the elemental principles (including in some cases weights of zero), it is unusual for underlying factors to influence judgment in qualitatively different directions.

Our findings regarding people's judgments of need accord with evidence supporting such multi-criteria theories (e.g., Frohlich and Oppenheimer, 1992; Scott et al., 2001; Konow, 2001, 2003). Differences between our work and this broader body of evidence, however, suggest fruitful avenues for future research. Konow's integrated approach, for instance, attempts to reconcile seemingly contradictory judgments of justice and fairness across decision contexts, but does not emphasize heterogeneous patterns of judgment within the population for a given context. In contrast, our work documents heterogeneous patterns of judgments regarding need within a single context, but does not explore heterogeneity across contexts. Future research into judgments of health care need should investigate different contexts (e.g., different health conditions) to test the robustness of the four patterns we identify, whether wholly new weighting patterns emerge in different contexts, and whether inclusion of a broader set of observable individual characteristics might help identify an association between participant characteristics and patterns of judgment. Analogously, future work on judgments of fairness in non-health-care domains should investigate heterogeneity of patterns of judgment within the population for a given decision context.

Our findings caution against the use of general phrases or principles, such as "allocation according to need," when eliciting the views of the public to inform resource allocation. Previous research (e.g., Hurley et al., 2011) documents important framing effects when eliciting public judgments about resource allocation whereby judgments differ importantly when allocation principles are described verbally (e.g., allocate according to need, allocate to maximize health, allocate to equalize health outcomes) compared to when the quantitative allocations implied by each principle in a given situation are presented. Further, when presented with both types of information, judgments correspond more closely to those made when given the quantitative information only rather than the verbal principles only. This study documents further that the public may interpret need quite differently than many health analysts. The same may hold for other important concepts, such as "benefit". To understand what the public thinks, studies must probe underneath high-level labels such as "need" to focus on the generic elements of an allocation problem.

Lastly, the latent-class, rank-ordered conditional logit estimator used in our analysis substantially outperformed the more commonly used conditional logit specification and enabled us to gain a much richer understanding of people's choices. This estimator has potentially widespread application in the empirical social choice studies that elicit individuals' rankings among a set of choice alternatives.

References

Abelson, J., J. Lomas, J. Eyles, S. Birch, and G. Veenstra (1995). Does the Community Want Devolved Authority? Results of Deliberative Polling in Ontario. *Canadian Medical Association Journal 153*, 403–412.

Ahlert, M., K. Funke, and L. Schwettmann (2013). Thresholds, Productivity, and Context: An Experimental Study of Determinants of Distributive Behaviour. *Social Choice and Welfare 40*(4), 957–984.

Allison, P. D. and N. Christakis (1994). Logit models for sets of ranked items. *Sociological Methodology 24*, 199–228.

Boulding, K. E. (1966). The Concept of Need for Health Services. *The Milbank Quarterly 44*(4), 202–223.

Braybrooke, D. (1987). *Meeting Needs.* Princeton, N.J.: Princeton University Press.

Cappelen, A. W., K. Moene, E. Sorensen, and B. Tungodden (2013). Needs vesus Entitlements– An International Fairness Experiment. *Journal of the European Economic Association 11*(3), 574–598.

Cuadras-Morato, X., J. L. Pinto-Prades, and J. M. Abellan-Perinoan (2001). Equity Considerations in Health Care: The Relevance of Claims. *Health Economics 10*, 187–205.

Culyer, A. J. (1995). Need: The Idea Won't Do - But We Still Need It. *Social Science and Medicine 40*(6), 727–730.

Culyer, A. J. and A. Wagstaff (1993). Equity and Equality in Health and Health Care. *Journal of Health Economics 12*(4), 431–457.

Daniels, N. (1985). *Just Health Care.* Cambridge: Cambridge University Pres.

Daniels, N. (2008). *Just Health: Meeting Health Needs Fairly.* New York, NY: Cambridge University Press.

Deutsch, M. (1985). *Distributive Justice.* Binghamton, N.Y.: Vail-Ballou Press.

Dillman, D. A. (2009). Response Rate and Measurement Differences in Mixed Mode Surveys Using Mail, Telephone, Interactive Voice Response and the Internet. *Social Science Research 38*(1), 1–18.

Evans, R. G. (1984). *Strained Mercy: The Economics of Canadian Health Care.* Toronto: Buttersworth.

Frohlich, N. and J. A. Oppenheimer (1992). *Choosing Justice: An Experimental Approach to Ethical Theory.* Berkeley, CA: University of California Press.

Gaertner, W. and E. Schokkaert (2012). *Empirical Social Choice.* Cambridge: Cambridge University Press.

Gajic, A., D. Cameron, and J. Hurley (2012). Cost-effectiveness of Cash versus Lottery Incentives for a Web-based, Stated Preference Survey. *European Journal of Health Economics 13*(6), 789–99.

Greene, W. and D. Hensher (2003). A Latent Class Model for Discrete Choice Analysis: Contrasts with Mixed Logit. *Transportation Research, Part B: Methodological 37*(8), 681–698.

Gu, Y., E. Lanscar, P. Ghijben, J. Butler, and C. Donaldson (2015). Attributes and Weights in Health Care Priority Setting: A Systematic Review of What Counts and to What Extent. *Social Science and Medicine 146*, 41–52.

Hasman, A., T. Hope, and L. Osterdal (2006). Health Care Need: Three Interpretations. *Journal of Applied Philosphy 23*(2), 145–156.

Herlitz, A. and D. Horan (2016). Measuring Need for Priority Setting in Health Care Planning and Policy. *Social Science and Medicine 157*, 96–102.

Hope, T., L. Osterdal, and A. Hasman (2010). An Inquiry into the Principles of Needs-based Allocation of Health Care. *Bioethics 24*(9), 470–480.

Hurley, J. (2000). An Overview of the Normative Economics of the Health Sector. In A. J. Culyer and J. P. Newhouse (Eds.), *Handbook of Health Economics*, pp. 55–118. Amsterdam: Elsevier Science B.V.

Hurley, J., N. Buckley, K. Cuff, M. Giacomini, and D. Cameron (2011). Judgments Regarding the Fair Division of Goods: The Impact of Verbal versus Quantitative Descriptions of Alternative Principles. *Social Choice and Welfare 37*(2), 341–372.

Juth, N. (2015). Challenges for Principles of Need in Health Care. *Health Care Analysis 23*, 73–87.

Kahneman, D. and C. Varey (1991). Notes on the Psychology of Utility. In J. Elster and J. Roemer (Eds.), *Interpersonal Comparisons of Well-Being*, pp. 127–163. Cambridge: Cambridge University Press.

Konow, J. (2001). Fair and Square: The Four Sides of Distributive Justice. *Journal of Economic Behavior and Organization 46*(2), 137–64.

Konow, J. (2003). Which is the Fairest of All? A Positive Analysis of Justice Theories. *Journal of Economic Literature 41*(4), 1188–1239.

Kuhfeld, W. (1997). *Efficient Experimental Design Using Computerized Search.* Sequim, WA: SAS Institute, Sawtooth Software Research Paper Series.

Kuhfeld, W. F. (2005). *Marketing Research Methods in SAS.* Cary, North Carolina: SAS Institute, Inc.

Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis.* New York: Dover.

Manfreda, L., M. Bosnjak, J. Hass, and V. Vehovar (2008). Web Surveys versus Other Survey Modes: A Meta-analysis Comparing Response Rates. *International Journal of Market Research 50*, 79–104.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. Frontiers in Econometrics, pp. 105–142. New York: Academic Press.

Mentzakis, E., P. Stefanowska, and J. Hurley (2011). A Discrete Choice Experiment Investigating Preferences for Funding Drugs used to Treat Orphan Diseases. *Health Economics, Policy and Law 6*(3), 405–433.

Miller, D. (1976). *Social Justice.* Oxford, UK: Clarendon Press.

Robertson, A. (1998). Critical Reflections on the Politics of Need: Implications for Public Health. *Social Science and Medicine 47*(10), 1419–1430.

Schwettmann, L. (2009). *Trading Off Competing Allocation Principles: Theoretical Approaches and Empirical Investigations.* New York: Peter Lang GmbH.

Scott, J. T., R. Matland, P. Michelbach, and B. Bornstein (2001). Just Desserts: An Experimental Study of Distributive Justice Norms. *American Journal of Political Science 45*(4), 749–767.

Shah, K. K. (2009). Severity of Illness and Priority-setting in Health Care: A Review of the Literature. *Health Policy 93*, 77–84.

Skitka, L. J. and P. E. Tetlock (1992). Allocating Scarce Resources: A Contingency Model of Distributive Justice. *Journal of Experimental Social Psychology 28*(6), 491–522.

Smith, R. (2007). Use, Option and Externality Values: Are Contingent Valuation Studies in Health Care Mis-specified? *Health Economics 16*(8), 861–869.

Swait, J. and W. Adamowicz (2001). The influence of task complexity on consumer choice: A latent class model of decision strategy switching. *Journal of Consumer Research 28*(1), 135–148.

Thokala, D. A. (2012). Multiple Criteria Decision Analysis for Health Technology Assessment. *Value in Health 15*, 1172–1181.

van Doorslaer, E., A. Wagstaff, and F. Rutten (1993). *Equity in the Finance and Delivery of Health Care: An International Perspective*. Oxford: Oxford University Press.

van Exel, J., R. Baker, H. Mason, C. Donaldson, and W. Brouwer (2015). Public Views on Principles for Health Care Priority Setting: Findings of a European Cross-Country Study using Q Methodology. *Social Science and Medicine 126*, 128–137.

Whitty, J., E. Lanscar, K. Rixon, X. Golenko, and J. Ratcliffe (2014). A Systematic Review of Stated Preference Studies Reporting Public Preferences for Health Care Priority Setting. *Patient 7*, 365–386.

Williams, A. (1978). Need - An Economic Exegesis. In A. Culyer and K. Wright (Eds.), *Economic Aspects of Health Services*, pp. 32–45. London: Martin Robertson.

Yaari, M. E. and M. Bar-Hillel (1984). On Dividing Justly. *Social Choice and Welfare 1*, 1–24.

Zwerina, K., J. Huber, and W. F. Kuhfeld (1996). *A General Method for Constructing Efficient Choice Designs*. SAS Technical Papers: TS-722E, SAS Inc.: Cary, North Carolina.

Table 1: Descriptive Statistics

	Full Sample (N=349)		Consistent Sample (N=212)		CCHS	
Variable	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Age (minimum = 18; maximum = 88)	53.5	13.91	53.9	13.68	44.7	19.27
Male (1 = male; 0 = female)	0.60	0.49	0.62	0.49	0.49	0.50
Married (1 = married; 0 = non-married)	0.73	0.45	0.72	0.45	0.50	0.50
SAHS (1 = E,VG; 0 = G,F,P)*	0.55	0.50	0.58	0.49	0.59	0.49
Education (1 = post-secondary; 0 = other)	0.78	0.41	0.82	0.38	0.74	0.44
Employed (1 = employed full-time; 0 = other)	0.57	0.50	0.56	0.50	0.48	0.50
Own House (1 = own home; 0 = rent)	0.89	0.31	0.87	0.33	0.72	0.45
Household Income						
< $20,000	0.023	−−	0.02	−−	0.072	−−
$20,000 to $49,999	0.180	−−	0.16	−−	0.264	−−
$50,000 to $99,999	0.335	−−	0.32	−−	0.342	−−
>$100,000	0.255	−−	0.29	−−	0.317	−−
Did not report	0.206	−−	0.21	−−	0.004	−−
Ever employed in health care sector	0.21	0.41	0.23	0.42	−−	−−

*E = Excellent, VG = Very Good, G = Good, F = Fair, P = Poor

Table 2: Latent-Class, Rank-Ordered Logit, Consistent Sample, No Interaction Terms, 4 Classes

	Class 1	Class 2	Class 3	Class 4
Mean Probability Class Assignment	0.378	0.189	0.337	0.096
Baseline Health				
Pain Free 4 hours	−2.968**	−2.093**	−1.155**	−0.013
Pain Free 8 hours	−4.764**	−3.828**	−2.174**	−0.066
Pain Free 12 hours	−7.316**	−5.335**	−3.947**	−0.061
Ability-to-Benefit				
4 hours relief	2.376**	−1.274**	4.022**	−1.633**
12 hours relief	0.976**	−0.605**	1.958**	−1.183**
16 hours relief	0.353**	−0.248	1.231**	−0.791**
Pills Required to Exhaust Benefit				
2 pills	−1.372**	1.053**	−1.825**	0.125
6 pills	−0.474**	0.637**	−0.725**	0.053
8 pills	−0.475**	0.401**	−0.618**	0.036
Observations	10,173			
Individuals	212			
Log-L	−3855.68			

Reference categories are: 0 hours pain-free per day; 24 hours of pain relief possible; and 12 pills

** $p < 0.05$; * $= 0.05 \leq p < 0.10$

Table 3: Latent-Class, Rank-Ordered Logit, Consistent Sample, Ratio ATB/RREB

	Class 1	Class 2	Class 3	Class 4
Mean Probability Class Assignment	0.390	0.183	0.355	0.072
Baseline Health				
Pain Free 4 hours	−2.769**	−1.985**	−0.769**	−0.244
Pain Free 8 hours	−4.396**	−3.638**	−1.300**	−0.419**
Pain Free 12 hours	−7.253**	−5.076**	−2.816**	−0.150
Hours of Pain Relief per Pill (ATB/Pills)				
Hours per pill = 1 to 2	−0.876**	0.716**	−1.259**	0.504**
Hours per pill = 2 to 4	−1.902**	1.260**	−2.480**	1.583**
Hours per pill > 4	−2.206**	1.601**	−2.590**	1.386**
Observations	10,173			
Individuals	212			
Log-L	−3928.90			

Reference categories are: 0 hours pain-free per day; Hours per pill < 1

** $p < 0.05$; * = $0.05 \leq p < 0.10$

Figure 1: **Example of a Choice Scenario Presented to Participants**

Question 1	Individual A	Individual B	Individual C
Hours each day free of pain with no treatment	0 hours	0 hours	0 hours
Additional hours of pain relief possible from medication	4 hours	12 hours	24 hours
Pain-relief pills required to obtain the maximum possible hours of pain relief from medication	6 pills	2 pills	8 pills

Please select the individual who you believe has the GREATEST need among individuals A, B and C:

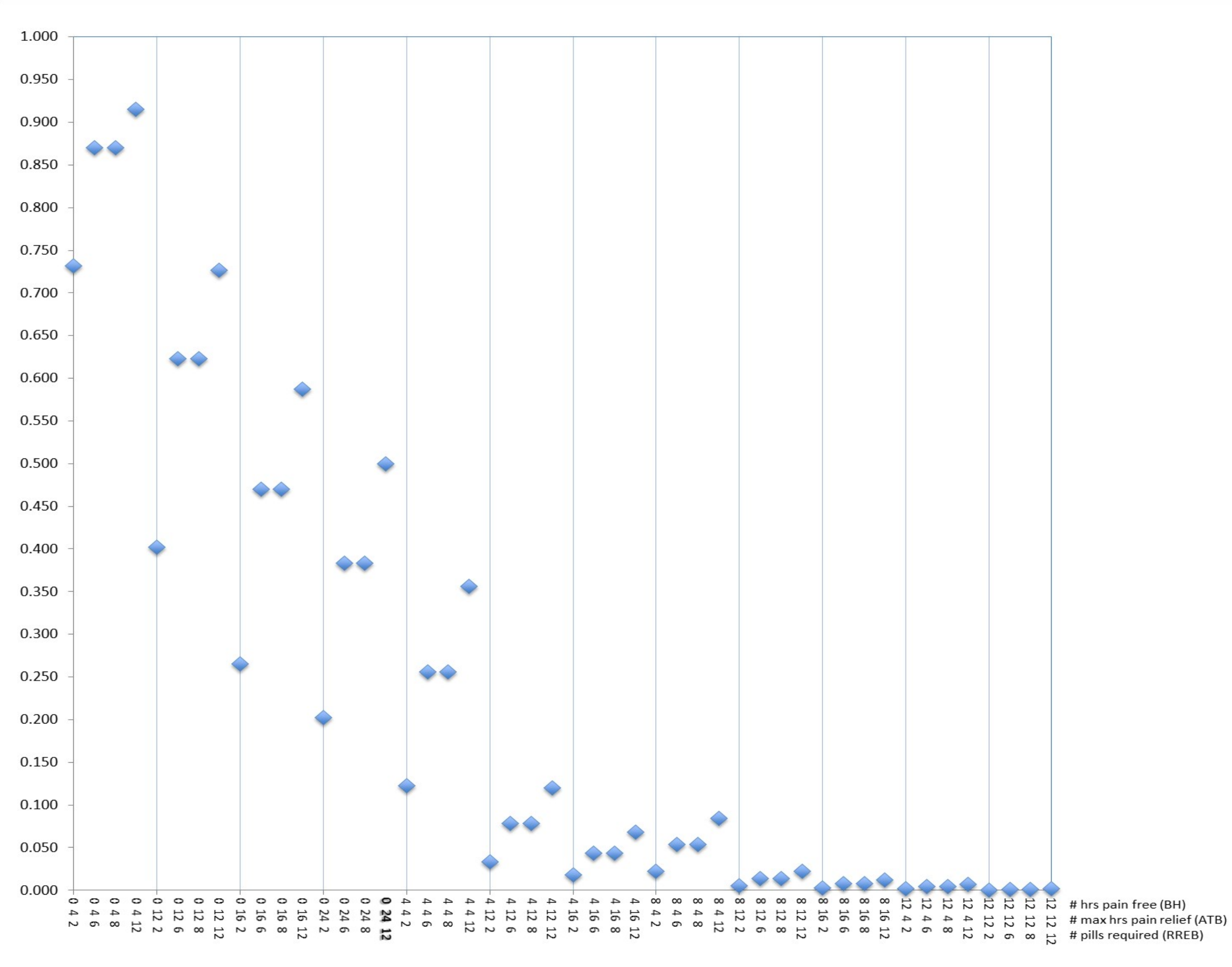
O Individual A   O Individual B   O Individual C

Please select the individual who you believe has the LEAST need among individuals A, B and C:

O Individual A   O Individual B   O Individual C

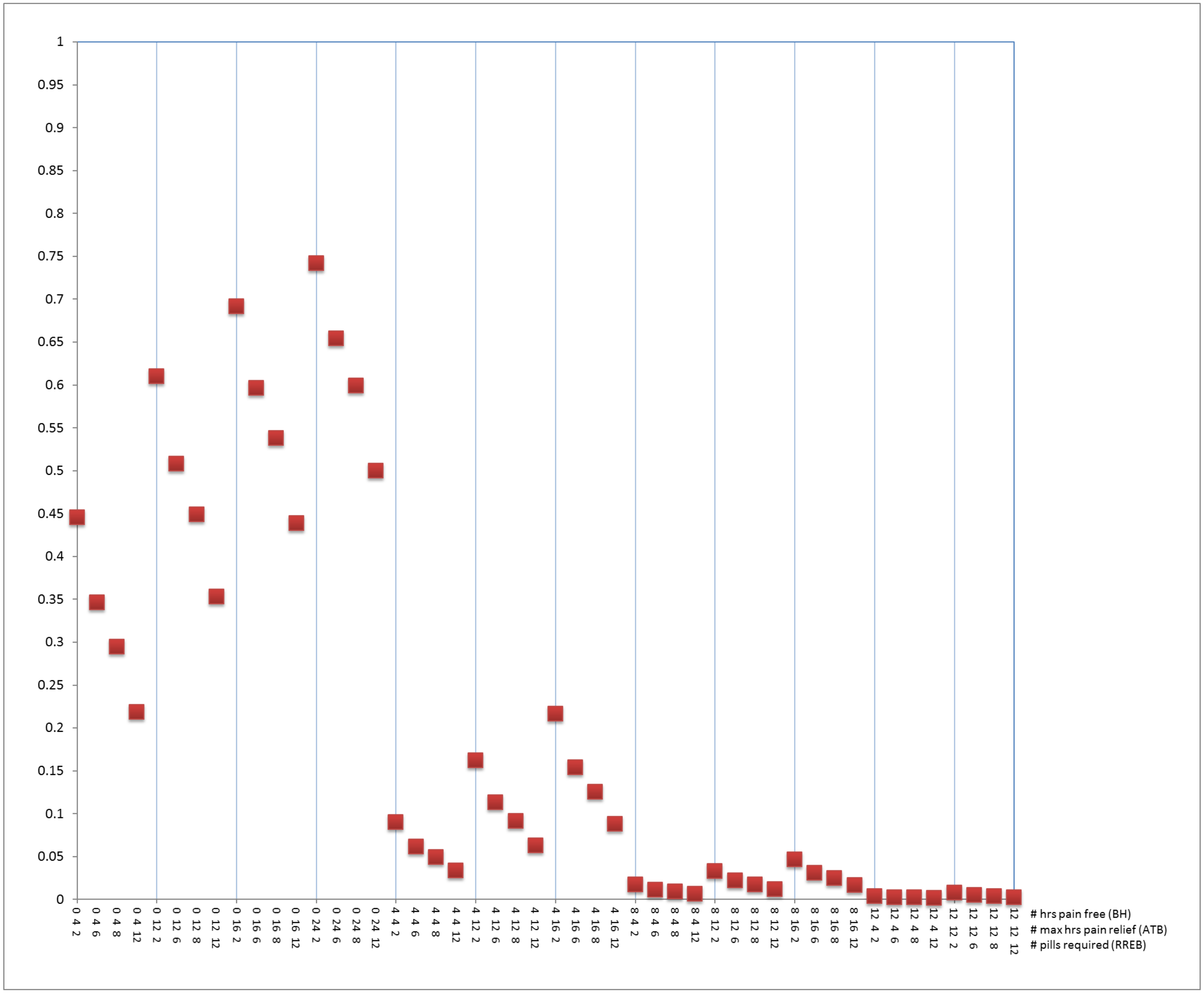
This section of the survey was then followed by 19 scenarios formatted identically to this. Only one scenario appeared on screen at a time, automatically followed by the next scenario upon response submission. Subjects were unable to go back to revise previous scenario responses after submitting. Note that if subjects submitted the same individual as having both the greatest and least need, a warning would appear: "You have chosen the same individual for both questions. Please make another selection" and subjects were then free to re-select either or both responses.

Figure 2: Class 1: Probability a Given Combination is Ranked as Greater Need than the Reference Combination



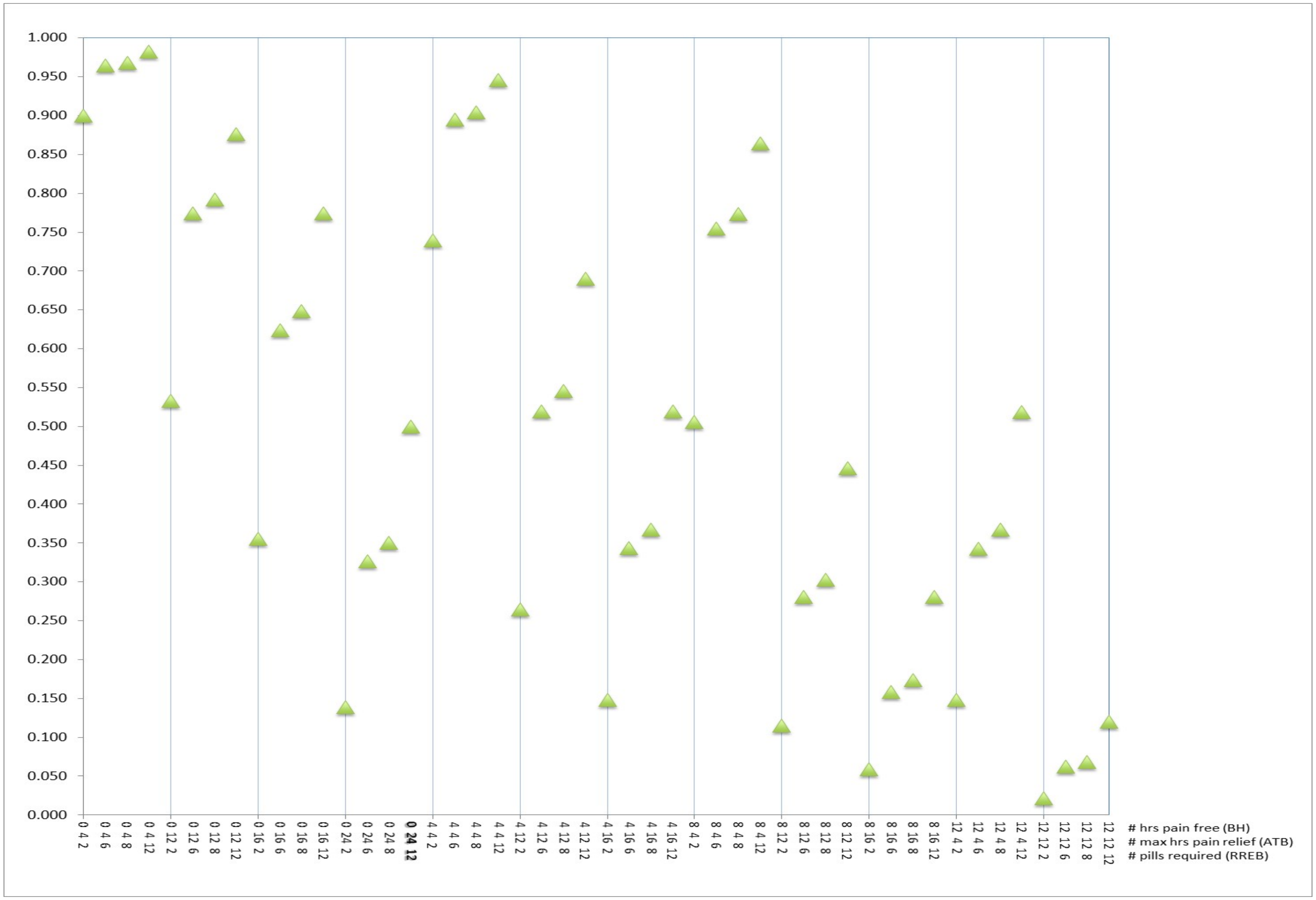
Notes: Reference combination: (Pain Free = 0, Pain Relief = 24 and Pills = 12). All probabilities are conditional on class.

Figure 3: Class 2: Probability a Given Combination is Ranked as Greater Need than the Reference Combination



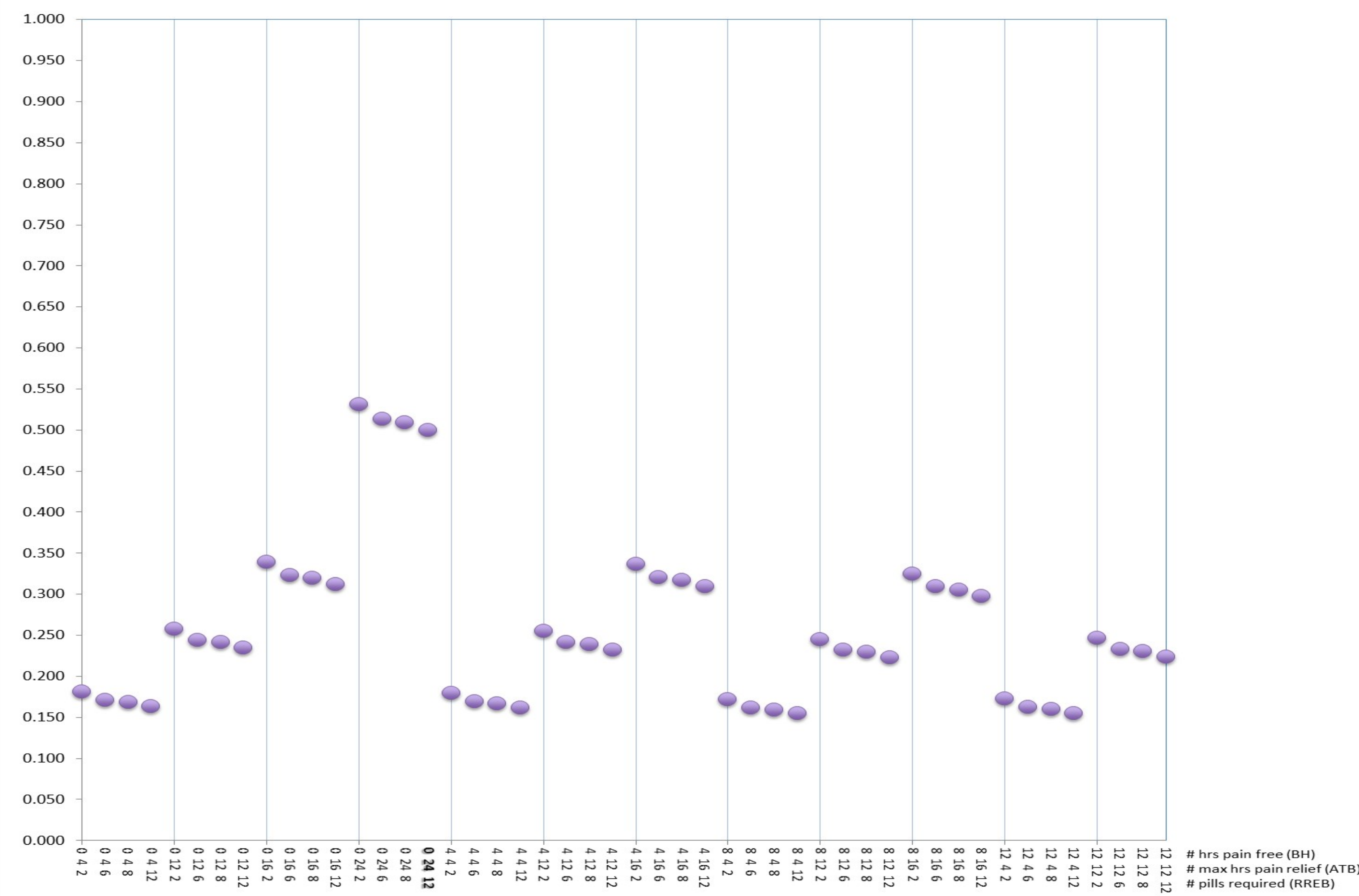
Notes: Reference combination: (Pain Free = 0, Pain Relief = 24 and Pills = 12). All probabilities are conditional on class.

Figure 4: Class 3: Probability a Given Combination is Ranked as Greater Need than the Reference Combination



Notes: Reference combination: (Pain Free = 0, Pain Relief = 24 and Pills = 12). All probabilities are conditional on class.

Figure 5: Class 4: Probability a Given Combination is Ranked as Greater Need than the Reference Combination



Notes: Reference combination: (Pain Free = 0, Pain Relief = 24 and Pills = 12). All probabilities are conditional on class.

Appendix 1: Simulations to Test Estimator Performance

The data generating process (DGP) for the simulations is based on a D-efficient experimental design with a two- and a three-level attribute, X and Z respectively, and six choice sets per individual. While simple, this design loses no generality over larger designs but requires less computational power. The GDPs tested are complex enough to incorporate all aspects of the latent-class framework that approximate well our estimation situation. We simulate models with two and three latent-classes. For the specification with two-classes, utilities for each class are given by

$$U_{Class_1} = \beta_{11}X + \beta_{12}Z_1 + \beta_{13}Z_2 + u_1 \qquad \text{and} \qquad U_{Class_2} = \beta_{21}X + \beta_{22}Z_1 + \beta_{23}Z_2 + u_2 \tag{6}$$

where $u_1, u_2 \sim EV1$. Assignment to class follows a multinomial logit[9] and is a function of an individual-specific variable k ($\sim N(0,1)$)

$$\pi_{Class_1} = \Lambda(\gamma_{11} + \gamma_{12}k) \tag{7}$$

For identification, the class probability parameters of Class 2 are normalized to zero and $\pi_{Class_2} = 1 - \pi_{Class_1}$. For the simulation with three latent-classes, we assume the same experimental design but the DGP incorporates a third latent class with utility

$$U_{Class_3} = \beta_{31}X + \beta_{32}Z_1 + \beta_{33}Z_2 + u_3, \quad \text{where } u_3 \sim EV1 \tag{8}$$

As before, class probabilities condition on k with

$$\begin{aligned} \pi_{Class_1} &= \frac{e^{(\gamma_{11}+\gamma_{12}*k)}}{e^{(\gamma_{11}+\gamma_{12}*k)} + e^{(\gamma_{21}+\gamma_{22}*k)} + 1} \\ \pi_{Class_2} &= \frac{e^{(\gamma_{21}+\gamma_{22}*k)}}{e^{(\gamma_{11}+\gamma_{12}*k)} + e^{(\gamma_{21}+\gamma_{22}*k)} + 1} \\ \pi_{Class_3} &= 1 - \pi_{Class_1} - \pi_{Class_2} \end{aligned} \tag{9}$$

All simulation experiments are based on 300 replications of 1000 individuals.

The results of the simulation exercise are given in Table 4 below. For each DGP we test a specification where class probability is based only on a constant and one where it conditions on an individual-specific variable k. Overall, the model performs well, correctly identifying signs and magnitudes of preference parameters (βs) and unobserved heterogeneity (γs) in both the two- and three-class models irrespective of conditioning on class-assignment covariates.

[9] With two classes this reduces to the standard logit model, as the parameters of Class 2 are normalized to zero.

Table 4: Latent class rank-ordered conditional logit simulation results

		Estimated Values			
		2-Classes		2-Classes with k	
		Mean	95% CI	Mean	95% CI
V_1	$\beta_{11} = 2$	2.003	(1.998, 2.007)	2.003	(1.998, 2.007)
	$\beta_{12} = 2$	2.002	(1.997, 2.008)	2.002	(1.997, 2.008)
	$\beta_{13} = 2$	2.002	(1.997, 2.007)	2.002	(1.997, 2.007)
V_2	$\beta_{21} = -2$	-2.000	(-2.006, -1.994)	-2.000	(-2.006, -1.994)
	$\beta_{22} = -2$	-2.006	(-2.013, -1.999)	-2.006	(-2.013, -1.999)
	$\beta_{23} = -2$	-2.009	(-2.015, -2.003)	-2.009	(-2.015, -2.003)
π_1	$\gamma_{11} = 0.5$	0.471	(0.464, 0.478)	0.500	(0.493, 0.507)
	$\gamma_{12} = 0.5$	-	-	0.495	(0.487, 0.503)
Class shares	$c_1 = 0.62$	0.615	(0.614, 0.617)	0.615	(0.614, 0.617)
	$c_2 = 0.38$	0.385	(0.383, 0.386)	0.385	(0.383, 0.386)
		3-Classes		3-Classes with k	
		Mean	95% CI	Mean	95% CI
V_1	$\beta_{11} = 2$	2.021	(2.010, 2.032)	2.019	(2.008, 2.030)
	$\beta_{12} = 2$	2.014	(2.003, 2.025)	2.014	(2.003, 2.025)
	$\beta_{13} = 2$	2.012	(2.001, 2.023)	2.011	(2.000, 2.023)
V_2	$\beta_{21} = -2$	-2.008	(-2.018, -1.997)	-2.007	(-2.018, -1.997)
	$\beta_{22} = -2$	-2.005	(-2.017, -1.993)	-2.005	(-2.017, -1.993)
	$\beta_{23} = -2$	-1.995	(-2.007, -1.984)	-1.995	(-2.006, -1.983)
V_3	$\beta_{31} = 1$	0.994	(0.981, 1.006)	0.995	(0.983, 1.008)
	$\beta_{32} = 1$	1.003	(0.989, 1.017)	1.003	(0.990, 1.017)
	$\beta_{33} = 1$	1.005	(0.992, 1.018)	1.006	(0.993, 1.019)
π_1	$\gamma_{11} = 0.5$	0.523	(0.418, 0.627)	0.468	(0.425, 0.511)
	$\gamma_{12} = 0.5$	-	-	0.533	(0.503, 0.564)
π_2	$\gamma_{21} = -0.5$	-0.467	(-0.567, -0.366)	-0.507	(-0.538, -0.476)
	$\gamma_{22} = -0.5$	-	-	-0.491	(-0.522, -0.459)
Class shares	$c_1 = 0.58$	0.578	(0.570, 0.586)	0.578	(0.571, 0.586)
	$c_2 = 0.13$	0.138	(0.137, 0.139)	0.138	(0.137, 0.139)
	$c_3 = 0.29$	0.284	(0.276, 0.292)	0.284	(0.276, 0.291)

Appendix 2: The Discrete-Choice Survey Instrument

HEALTH CARE NEEDS SURVEY [1]

Introduction

Allocating Health Care Resources According to Need

Canada, like many other countries around the world, has designed its health care system to provide health care services to individuals based on their need for care. However, it can be difficult for both health care providers and health system managers to assess people's need for care. Evaluating who has the greater need among those seeking care can be especially difficult. There is currently no universal agreement on what is most important in determining an individual's need for care.

The purpose of this short survey is to gain an understanding of your views regarding people's need for health care. In the survey we present you with a series of scenarios that describe three individuals who suffer from chronic pain. The level of pain experienced by the three individuals is identical, and is sufficient to keep them from participating in many of their normal daily activities. The individuals differ, however, in the amount of time that they are in pain each day when they receive no treatment, and in the hours of pain relief they can obtain from pain medication. Even with medication, some individuals are not able to obtain complete relief from pain. For each scenario, we are interested in your views on which of the three individuals has the greatest need for pain-relief pills, and which individual has the least need. There are no right or wrong answers.

The pain-relief pills that are available have no negative side effects for any of the individuals and can safely be taken in the amounts considered in the survey. The pain-relief pills are available to the individuals at no cost.

Each scenario describes three characteristics of the individuals related to the pain they experience and the relief that they can get from medication. All other characteristics of the individuals that are not described (such as an individual's age, sex, income, marital status, and so forth) are identical across the three individuals. The individuals differ only with respect to the three characteristics listed. The three characteristics are described on the next screen.

[1] The survey was designed and administered online using Limesurvey. This Appendix version is identical in content to the online versions but omits various "next buttons" used to progress to the

Characteristics

Characteristic 1: Number of hours each day free of pain with no treatment

This characteristic describes the number of pain-free hours in a 24-hour day that an individual experiences if they take no pain medication. This characteristic can take on four different values: 0, 4, 8 or 12 hours. For example, if the value for an individual is 4 hours and that individual takes no pain medication, they will be in pain 20 hours each day and free of pain 4 hours each day.

Characteristic 2: Number of hours of pain relief possible from medication

Some individuals are not able to obtain complete pain relief, no matter how many pills they take. This characteristic describes the maximum number of hours of pain relief an individual can obtain by taking pain-relief pills. This pain relief is in addition to any pain-free hours they experience if they take no medication. This characteristic can take on four different values: 4, 12, 16 or 24 hours. For example, if the value of this characteristic is 12 hours, then the maximum number of hours of additional pain relief the individual can achieve by taking medication 12 hours per day.

Characteristic 3: Number of pain-relief pills required to obtain the maximum possible hours of pain relief from medication

Because of biological differences among the individuals, the effectiveness of the pills differs across individuals. This characteristic describes the number of pain-relief pills an individual needs each day to obtain their maximum possible number of hours of pain relief per day. This characteristic can take on four different values: 2, 6, 8 or 12 pills. For example, if the value for an individual is 6 pills, then the individual must take 6 pills per day to obtain the greatest number of hours of pain relief possible.

Example of a Description

To give a concrete example, consider the following individual whose values for the three characteristics are as follows:

Number of hours free of pain with no treatment: 8 hours
Number of hours of additional pain relief possible: 12 hours
Number of pills required for maximum benefit: 6 pills

These three pieces of information tell you that this individual has 8 hours free of pain each day even if they take no pills, that by taking pain-relief pills they can obtain an additional 12 hours of pain relief (for a total of 20 hours per day free from pain), and that they must take 6 pills per day to achieve the maximum possible relief from pain.

Example of a Scenario

Individuals A, B and C all suffer to varying degrees from chronic pain. A limited supply of pain-relief medication is available that relieves this chronic pain, but its effectiveness differs across individuals. The limited supply of pain-relief pills is not sufficient to provide complete pain relief to all three individuals. Based on the information provided in the table below, please indicate which individual you judge to have the *greatest* need, and which other individual you judge to have the *least* need.

Example	**Individual A**	**Individual B**	**Individual C**
Hours each day free of pain with no treatment	0 hours	4 hours	8 hours
Additional hours of pain relief possible from medication	4 hours	12 hours	12 hours
Pain-relief pills required to obtain the maximum possible hours of pain relief from medication	6 pills	2 pills	8 pills

Please select the individual who you believe has the GREATEST need among individuals A, B and C:

O Individual A O Individual B O Individual C

Please select the individual who you believe has the LEAST need among individuals A, B and C:

O Individual A O Individual B O Individual C

(This is an example, please do not answer)

Note that we have created reminder pop-up windows to help remind you of the characteristics. Simply click on any of the characteristic labels for helpful information.

On the following 20 screens, we present 20 different scenarios that are similar to this example. For each of the scenarios, please choose the most preferred option. Remember, there are no right or wrong answers.

Question 1	Individual A	Individual B	Individual C
Hours each day free of pain with no treatment	0 hours	0 hours	0 hours
Additional hours of pain relief possible from medication	4 hours	12 hours	24 hours
Pain-relief pills required to obtain the maximum possible hours of pain relief from medication	6 pills	2 pills	8 pills

Please select the individual who you believe has the GREATEST need among individuals A, B and C:

O Individual A O Individual B O Individual C

Please select the individual who you believe has the LEAST need among individuals A, B and C:

O Individual A O Individual B O Individual C

This section of the survey was then followed by 19 scenarios formatted identically to this. Only one scenario appeared on screen at a time, automatically followed by the next scenario upon response submission. Subjects were unable to go back to revise previous scenario responses after submitting. Note that if subjects submitted the same individual as having both the greatest and least need, a warning would appear: "You have chosen the same individual for both questions. Please make another selection" and subjects were then free to re-select either or both responses.

Demographic Questions

In this last part of the survey we gather some basic information about you. The information from the following questions will not be used to identify you; it will assist in the analysis. Please select one answer for each of the questions.

What is your sex?

O Female
O Male
O Prefer not to respond

What is your year of birth? _________

In general, compared to individuals of your own age, would you say your health is (excellent, very good, good, fair, poor)?

O Excellent
O Very Good
O Good
O Fair
O Poor
O Prefer not to respond

What is your marital status?

O Single (never legally married)
O Married or Common-Law
O Divorced
O Widowed
O Prefer not to respond

Are you currently employed?

O Part-time employment
O Full-time employment
O Not Employed
O Retired
O Prefer not to respond

Do you own or rent your house / apartment?

O Own
O Rent
O Other
O Prefer not to respond

What is the highest level of education you have attained?

O Less than secondary school
O Secondary school graduate
O Post-secondary graduate (i.e., college, apprentice, trade diploma or certificate)
O University graduate
O Prefer not to respond

What is your best estimate of your total household income over the past 12 months?

O No income
O Less than $20,000
O $20,000 to $49,999
O $50,000 to $99,999
O $100,000 or more
O Prefer not to say

Have you ever worked in the health care sector?

O Yes
O No
O Prefer not to respond

Click Here to Submit Your Survey.

Thank you for completing this survey.

Appendix 3: Latent-Class, Rank-Order Logit Model Estimated over the Full Sample, No Interaction Terms, 4 Classes

	Class 1	Class 2	Class 3	Class 4
Mean Probability Class Assignment	0.292	0.204	0.219	0.285
Baseline Health				
Pain Free 4 hours	-2.918^{**}	-1.212^{**}	-1.388^{**}	-0.311^{**}
Pain Free 8 hours	-4.611^{**}	-2.222^{**}	-2.452^{**}	-0.550^{**}
Pain Free 12 hours	-6.915^{**}	-3.050^{**}	-4.390^{**}	-0.886^{**}
Ability-to-Benefit				
4 hours relief	2.008^{**}	-2.111^{**}	4.666^{**}	0.483^{**}
12 hours relief	0.870^{**}	-1.218^{**}	2.182^{**}	0.157
16 hours relief	0.275^{*}	-0.663^{**}	1.340^{**}	0.053
Pills Required to Exhaust Benefit				
2 pills	-1.287^{**}	1.145^{**}	-1.977^{**}	-1.029^{**}
6 pills	-0.496^{**}	0.677^{**}	-0.693^{**}	-0.574^{**}
8 pills	-0.428^{**}	0.446^{**}	-0.612^{**}	-0.318^{**}
Observations	16,743			
Individuals	349			
Log-L	-7114.28			

Reference categories are: 0 hours pain-free per day; 24 hours of pain relief possible; and 12 pills
** $p < 0.05$; * $= 0.05 \leq p < 0.10$

Appendix 4: Latent-Class, Rank-Order Logit Model Estimated over the Consistent Sample, Two-way Interaction Terms, 4 Classes

	Class 1	Class 2	Class 3	Class 4
Mean Probability Class Assignment	0.576	0.213	0.137	0.074
Baseline Health				
Pain Free 4 hours	−1.690**	−2.738**	0.535	−0.012
Pain Free 8 hours	−3.855**	−2.699**	−0.417	−0.886
Pain Free 12 hours	−5.890**	−6.316**	−2.132**	−0.115
Abilit-to-Benefit				
4 hours relief	3.547**	−1.058**	2.412**	−2.603**
12 hours relief	1.818**	−0.555**	1.337**	−1.847**
16 hours relief	1.234*	−0.201	0.368	−1.300**
Pills Required to Exhaust Benefit				
2 pills	−1.440**	0.641**	−2.241**	0.511
6 pills	−0.748**	0.338	−1.025**	−0.023
8 pills	−0.062**	−0.147	−0.238	−0.164
Interactions: BH x ATB				
pf4atb4	0.824**	−0.060	−0.565	−0.331
pf4atb12	0.282	−0.041	−0.475	−0.432
pf8atb4	0.095	0.323	−1.058**	−0.242
pf8atb12	−0.022	0.095	−0.612	−0.120
pf12atb4	−0.113	0.449	0.256	−0.812
Interactions: BH x RREB				
pf4pill2	0.706**	0.317	−0.827	−0.026
pf4pill6	−0.520*	0.499	−1.163**	0.316
pf4pill8	−1.716**	1.046**	−1.404*	0.794
pf8pill2	0.078	−1.843**	−0.426	1.206
pf8pill6	0.850**	−1.744*	−0.240	1.394
pf8pill8	0.419	−1.543*	−0.191	0.808
pf12pill2	1.237**	0.617	−0.137	0.397
pf12pill6	0.916**	−0.020	0.058	0.629
pf12pill8	0.505	1.624**	−0.222	−0.303
Observations	10,173			
Individuals	212			
Log-L	−3702.59			

Reference categories are: 0 hours pain-free per day; 24 hours of pain relief possible; and 12 pills
** $p < 0.05$; * = $0.05 \leq p < 0.10$

Appendix 5: The Sub-study Debriefing Interview Instrument

NEED SURVEY Semi-Structured Interview Questions, version 2

1. How easy or difficult were the survey questions for you?
    a. Were any parts of the survey confusing? How?
    b. Did you encounter any [other] problems? Which?

    Prompts: If possible, prompt for what they would change, and how, beyond the complaint

2. Some questions asked you which person had the greatest need.
    a. What does "need" mean to you in the context of the survey scenarios? How did you think about it?
    b. What does "need" for health care mean to you, more generally? How do you think about need?

3. Now I'd like to go over a couple of the scenarios you saw, and would like you to tell me more about how you thought these through.

    c. [present Scenario 1]

    - Was this scenario difficult to understand in any way? How?
    - You indicated that individual XX has the greatest need and that this is because YYY.
        - How did you come to this choice?
        - Was this choice easy or difficult to make? How?
    - You indicated that individual XX has the least need and that this is because YYY.
        - How did you come to this choice?
        - Was this choice easy or difficult to make? How?
        -

    d. [present Scenario 2]

    - Was this scenario difficult to understand in any way? How?
    - You indicated that individual XX has the greatest need and that this is because YYY.
        - How did you come to this choice?
        - Was this choice easy or difficult to make? How?
    - You indicated that individual XX has the least need and that this is because YYY.
        - How did you come to this choice?
        - Was this choice easy or difficult to make? How?

4. I would now like to focus a bit on each of the characteristics of the individuals.

- *Hours free of pain with no medication*: How important was this for assessing need? Why was this important to you --- in what way does this represent a need ?

- *A person's ability to get pain relief from medication*: How important was this. Based on your earlier response, it seems that the greater is a person's ability to obtain pain relief, the greater [lesser] is their need? Why is that?

-
- Finally, the *number of pills required to obtain the maximum possible pain relief.* Based on your earlier response, it seem that the more pills a person requires, the greater (lesser) is their need? Again, why is that?

5. Did how you think about a person's need change at all as you worked through the survey? How?

6. Is there anything else you would like us to know about your experience taking this survey, or making these choices? Do you have any questions for me?

Scenario 1 (question 3 in full survey):

Individual	**A**	**B**	**C**
Hours each day free of pain with no treatment	4 hours	12 hours	4 hours
Additional hours of pain relief possible from medication	12 hours	4 hours	12 hours
Pain-relief pills required to obtain the maximum possible hours of pain relief from medication	2 pills	6 pills	8 pills

Please select the individual who you believe has the **GREATEST** need among individuals A, B and C:
Please explain why you believe this individual has the **GREATEST** need?
Please select the individual who you believe has the **LEAST** need among individuals A, B and C:
Please explain why you believe this individual has the **LEAST** need.

Scenario 2 (question 17 in full survey):

Individual	A	B	C
Hours each day free of pain with no treatment	8 hours	0 hours	12 hours
Additional hours of pain relief possible from medication	16 hours	4 hours	12 hours
Pain-relief pills required to obtain the maximum possible hours of pain relief from medication	8 pills	2 pills	6 pills

Please select the individual who you believe has the **GREATEST** need among individuals A, B and C:
Please explain why you believe this individual has the **GREATEST** need?
Please select the individual who you believe has the **LEAST** need among individuals A, B and C:
Please explain why you believe this individual has the **LEAST** need.

Scenario 2 *FOR PILOT (interview 1-5)* (question 8 in full survey):

Individual	A	B	C
Hours each day free of pain with no treatment	8 hours	0 hours	0 hours
Additional hours of pain relief possible from medication	12 hours	24 hours	16 hours
Pain-relief pills required to obtain the maximum possible hours of pain relief from medication	12 pills	2 pills	8 pills

Please select the individual who you believe has the **GREATEST** need among individuals A, B and C:
Please explain why you believe this individual has the **GREATEST** need?
Please select the individual who you believe has the **LEAST** need among individuals A, B and C:
Please explain why you believe this individual has the **LEAST** need.

NB: Questions 3 and 8 were used for the scenarios in the pilot interviews (1-5). We subsequently switched to Question 3 and 17

Appendix 6: Latent-Class, Rank-Order Logit Model, Consistent Sample, No Interactions, 4 Classes, Follow-up Study

	Class 1	Class 2	Class 3	Class 4
Mean Probability Class Assignment	0.343	0.372	0.114	0.171
Baseline Health				
Pain Free 4 hours	−2.522**	−3.311**	−1.696**	0.229
Pain Free 8 hours	−3.312**	−4.715**	−4.282**	−0.262
Pain Free 12 hours	−5.024**	−6.108**	−6.440**	−0.749*
Ability-to-Benefit				
4 hours relief	1.020**	−3.260**	6.091**	−3.773**
12 hours relief	−0.064	−1.494**	3.716**	−2.509**
16 hours relief	−0.490	−0.395	2.668**	−1.524**
Pills Required to Exhaust Benefit				
2 pills	−1.970**	1.744**	−1.736**	2.612**
6 pills	−1.215**	1.159**	−1.391**	1.826**
8 pills	−0.827**	0.160	−1.183**	1.017***
Observations	1,680			
Individuals	35			
Log-L	−568.030			

Reference categories are: 0 hours pain-free per day; 24 hours of pain relief possible; and 12 pills

** $p < 0.05$; * $= 0.05 \leq p < 0.10$