

Understanding the formation and influence of attitudes in patients' treatment choices for lower back pain: testing the benefits of a hybrid choice model approach

Mirja Elisabeth Kløjgaard

COHERE, University of Southern Denmark, mirja.kloejgaard@hotmail.com

Stephane Hess

Institute for Transport Studies, University of Leeds, s.hess@its.leeds.ac.uk

Abstract

A growing number of studies across different fields is making use of a new class of choice models, labelled variably as hybrid model structures or integrated choice and latent variable models, in incorporating the role of attitudes in decision making. To date, this technique has not been used in health economics. The present paper looks at the formation of such attitudes and their role in patients' treatment choices in the context of low back pain. We use data stated choice data collected from a sample of just under 300 respondents referred to a regional spine centre in Denmark. We show how the hybrid model structure is able to make a link between attitudinal questions and treatment choices, and also explains the variation of these attitudes across key socio-demographic groups. However, we also show how, despite their growing popularity, the key findings of the model, net of a greater insight into the drivers of attitudes and small gains in efficiency, are no different from standard approaches which remain much easier to apply. In the present application, we also show how only a small share of the heterogeneity can be linked to the attitudinal construct.

Keywords: lower back pain; hybrid choice models; latent variables; stated choice

Introduction

In the study of choices of treatments, discrete choice models are widely used and accepted. These models are used to explain behaviour of patients and HCPs (see de Bekker-Grob et al., 2012, for a review). Much research has been done concentrating on how to explain differences in utilities between decision makers, focusing on deterministic heterogeneity, e.g. through socio-demographic interactions in a Multinomial Logit Model (MNL) or random heterogeneity in a Mixed Multinomial Logit Model (MMNL) model, or a combination of the two in a Latent Class Models (LCM). An example of such a study is the work of Hole (2008).

When patients are in a situation where they have to reach a decision whether or not to have treatment for a condition, or what specific treatment to choose, different patients will make different choices. Some of this heterogeneity can be attributed to the severity of their condition, or past experience, but there is clear scope also for idiosyncratic preferences, and we believe, a key role for attitudes and perceptions. This could be especially the case in situations where considerable uncertainty exists and where even health care practitioners (HCP) are experts. One such example is the case of low back pain (LBP), where diagnosis and the choice of treatment (i.e. surgical versus non-surgical), is distorted by conflicting evidence and by no promise of recovery with either modality (Allen et al., 2009; van Tulder et al., 2002; Gibson, 2005). This has resulted in a remarkable variance in surgery rates within regions and counties (Irwin et al., 2005a,b; Bederman et al., 2010) and in discussions of the (cost)effectiveness and prioritisation of treatment, both politically and amongst HCPs (Balagué et al., 2012). MIRJA: I WOULD REMOVE THE NEXT TWO SENTENCES. THEN, IN THE LAST SENTENCE, COULD WE FOCUS MORE ON PATIENTS AND USE APPROPRIATE REFERENCES FOR SUCH WORK? When HCPs and patients have to make a difficult choice between treatment options, the principal-agent relationship is not made any easier when evidence is also unclear. In spite of, or potentially because of unclear decision contexts, patients as well as HCPs form perceptions relating to the available options based on additional factors not relating to clinical evidence. Studies have shown that HCPs treat patients differently in accordance with their own perception of diagnosis and that HCPs are heavily guided by their beliefs about back pain in general and about the individual patient, in their treatment recommendations (Coudeyre et al., 2006; Poiraudau et al., 2006; Houben et al., 2005, 2004; Corbett et al., 2009; Pincus et al., 2007; Balagué et al., 2012; Bishop et al., 2008;

Linton et al., 2002; Main et al., 2010)

In line with the above, this paper argues that, possibly particularly in scenarios where clinical evidence is limited or not clear cut, perceptions that the patient forms, either through past experience or through discussion with other patients and/or medical experts, will play a major role in shaping his or her decisions. Researchers outside of health economics have increasingly recognised that a large share of this heterogeneity could be linked to underlying values, perceptions, attitudes and beliefs (Ben-Akiva et al., 1999). In settings where we know from contextual evidence that perceptions are vital in decision making, such as in the treatment of low back pain, the inclusion of such factors in our models is arguably especially important.

Surveys in health economics routinely ask questions of respondents with a view to capturing information on underlying attitudes and perceptions. Directly including such responses in the specification of the utility function of a choice model may seem tempting and may well produce reasonable effects. However, as recognised in a growing body of research (Ben-Akiva et al., 1999; Bolduc and Alvarez-Daziano, 2010, see e.g.), this is theoretically inappropriate and could potentially lead to substantial bias in model outputs and an inability to use a model in forecasting. Two quite distinct factors are at play, relating to risk of endogeneity bias and measurement error. Firstly, responses to questions of an attitudinal nature are likely to be correlated with other unobserved factors which enter the model's error term. If such answers are included in the modelled component of utility, this thus creates a potential for endogeneity bias due to correlation between the modelled and random utility components. Secondly, it should be clear that the answers given to such questions are not direct measures of attitudes or perceptions, but merely a function of such underlying factors. Researchers in the fields of transport, marketing and environmental economics are increasingly acknowledging this by treating these psychological constructs as latent variables in their models. For some examples of applications across different fields, see Hess and Beharry-Borg (2012); Abou-Zeid et al. (2010); Daly et al. (2012); Daziano and Bolduc (2011).

To the best of our knowledge, researchers in health economics are yet to make use of such hybrid structures in their work. The aim of the present paper is to investigate their potential benefits in capturing the role that perceptions and attitudes may have in explaining treatment choices made by patients. This is done by using stated choice (SC) data collected at a large Danish Spine Clinic, where the survey explicitly explored the process by which perceptions and attitudes are formed, drawing

on their past experience and attitudes to treatment. In our empirical work, we analyse how these perceptions influence patients' choices of treatment. This is achieved in a joint model of the formation of perceptions and of the choices made in the survey, using state-of-the-art hybrid model structures. We contrast the findings from this model to structures allowing for simple random heterogeneity and show how, while the hybrid structures provide some further insights into the formation of attitudes, and some gains in efficiency, the overall results remain largely unaffected. This finding should serve as a reminder to academics and practitioners that this new type of model is not some *magic bullet* which will radically change results from choice models.

The remainder of this paper is organised as follows. The next section gives an overview of the modelling methodology concerning the integration of choice and latent variables in a hybrid model. This is followed by a description of the empirical data and a discussion of model specification for our specific case study. Next, we present and discuss the econometric results for both samples. Finally the paper discusses the findings and provides recommendations for future research.

Modelling methodology

This section gives a brief overview of hybrid model structures and their use in incorporating the role of attitudes or perceptions in choice models. For more extensive details, the reader is referred to Ben-Akiva et al. (1999, 2002a,b); Bolduc et al. (2005); Daly et al. (2012).

In a standard random utility model, we have that the utility of alternative i as faced by respondent n in choice task t is given by:

$$U_{n,i,t} = V_{n,i,t} + \varepsilon_{n,i,t} \tag{1}$$

where $V_{n,i,t}$ and $\varepsilon_{n,i,t}$ give the deterministic and random component of utility, respectively. In a traditional model, we would have that $V_{n,i,t} = f(x_{n,i,t}, z_n, \beta)$, i.e. the deterministic component of utility is given by a function of the attributes of the alternative, $x_{n,i,t}$, measured characteristics of the respondent z_n , and estimated model parameters β , also often referred to as tastes or sensitivities. In many cases, $f()$ will equate to a linear-in-attributes specification, but we allow for any degree of flexibility with our notation.

The vector z_n contains respondent characteristics such as income, age and gender. Imagine now a

situation where as part of a survey, an analyst also captures answers from a respondent to L questions about attitudes, perceptions and convictions. Let I_n be a vector grouping together these answers, which may take a variety of format, for example being continuous or ordinal in nature, or simple binary yes/no answers. The key reasoning for using hybrid structures is that the simple inclusion of I_n in the utility function $V_{n,i,t}$ is theoretically misguided and could lead to substantial problems in model results. In particular, any answers to attitudinal questions or questions about perceptions are not direct measures of such attitudes or perceptions, but only functions thereof, or *indicators*. The simple inclusion of I_n in $V_{n,i,t}$ could thus lead to problems with measurement error, where this is further compounded by the fact that the values in I_n are often captured on a ordinal scale. Secondly, there is likely to be correlation between the answers in I_n and other unobserved factors influencing the behaviour of respondent n - the fact that such factors are captured in $\varepsilon_{i,n,t}$ could thus lead to correlation between $V_{n,i,t}$ and $\varepsilon_{i,n,t}$ and a risk of endogeneity bias. Finally, when forecasting of decisions is of interest, as it often is especially in fields such as transport or marketing, then values of I_n would not be available in the forecast period, making forecasting impossible.

The approach taken to deal with these problems in hybrid models is to see I_n as a dependent variable rather than an explanatory variable. In particular, we hypothesise that the true underlying attitudes and perceptions of respondent n , described by a vector of K unobserved (or latent) variables α_n , are influencing the answers that a respondent gives to questions of an attitudinal or perceptual nature (i.e. I_n) while also driving the behaviour in the actual choice situations. To this extent, α_n is used in such models to explain both I_n and C_n , where the latter refers to the sequence of choices observed for respondent n . Figure 1 show the outline of such a hybrid model structure.

The latent variable α_n is by nature unobserved and a key component of it is given by a vector of K random disturbances, ξ_n , which are assumed to follow a Normal distribution with a mean of zero and a covariance matrix Ω_ξ , where we typically assume that the off-diagonal elements in Ω_ξ are all zero, i.e. the individual latent variables are uncorrelated. In addition to the random component, we also allow for deterministic effects in α_n , specifically through socio-demographic interactions, such that:

$$\alpha_n = g(z_n, \gamma) + \xi_n, \tag{2}$$

where ξ_n is already defined above, as is z_n , and where γ is a matrix of K rows (one per latent variable)

with sufficient columns to cover all elements in z_n . Once again, the specific functional form used for $g()$ is left to the analyst to decide, but generally, a linear specification will be used. Equation 2 gives the structural equation for the latent variables in the hybrid model.

As a next step, we use α_n to explain the values of the indicators of attitudes and perceptions for respondent n , specifically using:

$$I_n = \delta_{I_n} + h(\alpha_n, \zeta) + \psi_n, \quad (3)$$

where once again a decision needs to be made on the functional form of $h()$ and where δ_{I_n} is a vector of constants, ζ is a vector of estimated parameters and ψ_n is a random disturbance. Equation 3 gives the measurement model for the indicators. The specific assumption for the random distribution of ψ has an impact on the likelihood function for the observed values. For example, if an assumption is made that the elements in ψ_n are normally distributed, then the probability of the observed value for indicator $I_{n,l}$ would now be given by:

$$P_{I_{n,l}} = \frac{1}{\sigma_{I,l}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{I_{n,l} - \delta_{I,l} - \zeta_{l,k}\alpha_{n,k}}{\sigma_{I,l}}\right)^2} \quad (4)$$

In this specification, we assume a relationship between this indicator $I_{n,l}$ and latent variable $\alpha_{n,k}$, and where the same latent variable is potentially used for multiple indicators. The inclusion of the constant for indicator $I_{n,l}$, i.e. $\delta_{I,l}$ becomes superfluous if we centre all indicators on zero by subtracting their mean. In the above specification, a positive value for $\zeta_{l,k}$ would mean that as $\alpha_{n,k}$ increases, the likelihood of a higher value for $I_{n,l}$ would increase too. Various other approaches are possible, where e.g. Daly et al. (2012) advocate the use of an ordered logit approach for ordinal indicators. Independent of the approach used for individual indicators, including where a mix of approaches is used across them, we can now write the probability of the observed set of respondent provided answers as:

$$P_{I_n|\alpha_n, \zeta, \Omega_I} = \prod_{l=1}^L P_{I_{n,l}}, \quad (5)$$

where ζ is a vector of estimated parameters showing the impact of the various latent variables on the various indicators, and where Ω_I is a set of parameters relating to the specification of the measurement

model, for example standard deviations in the case of normal densities, or thresholds in the case of an ordered logit or probit specification.

The final component in the hybrid model is the choice model component, where the latent variables are now incorporated into the utility specification, such that $V_{n,i,t} = f(x_{n,i,t}, \alpha_n, z_n, \beta, \tau)$, where τ is a vector of parameters that measures the impact of the latent variable in the utility function. This could consist of interactions with alternative specific constants and/or marginal utility coefficients.

The estimation of this hybrid structure now entails a set of additional parameters in comparison with a basic model. In particular, we need to estimate the vector of parameters γ which link the latent variable to socio-demographic characteristics of the respondent, the set of parameters Ω_I which are the parameters used in the measurement model, and the vector of parameters τ which capture the impact of the latent variable in the utility functions of the choice model. There is also the vector of diagonal elements of Ω_ξ , i.e. the variances used in the structural equations for the latent variable. A normalisation of the scale of the latent variables is required. Two different normalisations exist, either putting a constraint on the vector τ , as done by Ben-Akiva et al. (1999), or on the variances of the latent variables, as done by Bolduc et al. (2005). As shown empirically by Daly et al. (2012), the two normalisations are formally equivalent. In addition to these various parameters, we also need to estimate the core choice model parameters β , either as point estimates or the parameters of their distribution in a random coefficients model¹.

The hybrid model is thus made of two key components, a choice model and a measurement model, both of them depending on α_n , and both components are estimated simultaneously², with final log-likelihood function given by:

$$LL(\Omega_\beta, \gamma, \tau, \zeta, \Omega_I) = \sum_{n=1}^N \ln \int_{\beta} \int_{\alpha} P_{C_n} P_{I_n} \phi(\xi) m(\beta | \Omega) d\beta d\alpha \quad (6)$$

where we use the Bolduc et al. (2005) normalisation, such that no elements of Ω_ξ need to be estimated. In Equation 6, P_{C_n} gives the likelihood of the observed sequence of T_n choices for respondent n , which will typically be given by a product of logit probabilities (allowing for random heterogeneity through

¹In our application, we rely on point estimates, i.e. an absence of additional random heterogeneity.

²Sequential estimation is also possible, but leads to a loss of efficiency.

the integration over β). In particular, we would have:

$$P_{C_n} = \prod_{t=1}^{T_n} \frac{e^{V_{i_n^*,t}^*}}{\sum_{j=1}^J e^{V_{n,j,t}}}, \quad (7)$$

where $i_{n,t}^*$ is the alternative chosen by respondent n in task t , and where as stated above, we have that $V_{n,i,t} = f(x_{n,i,t}, \alpha_n, z_n, \beta, \tau)$. Next, P_{I_n} gives the likelihood of the observed sequence of answers to the attitudinal questions, which is given by 5. Both P_{I_n} and P_{C_n} depend on the latent component α_n , while the P_{C_n} also depends on the randomly distributed β parameter. Integration of the product of P_{C_n} and P_{I_n} over the distribution of β and α is thus needed, and this explains the presence of the density function for the random component in α , i.e. $\phi(\xi)$ and the density function for β , i.e. $m(\beta | \Omega_\beta)$, in Equation 6. The latter is a function of an estimated vector of parameters Ω_β , while the parameters of the former have been normalised for identification (means to 0, variances to 1).

In practice, Equation 6 does not possess a closed form solution such that typically, simulation based estimation of the model is used, evaluating $P_{C_n} P_{I_n}$ at a large number of draws from β and α . Finally, when no random heterogeneity in β is accommodated in the model, the integration over β drops out from Equation 6, and we estimate a vector of point values for β , rather than the parameters of its distribution, Ω_β .

The contrast between this hybrid model and deterministic approaches employing the answers to attitudinal questions as explanatory variables in the utility function is that the hybrid model still makes use of these answers I_n but treats them as dependent variables in a measurement component of the joint model rather than as explanatory variables. The link between the two components is made through the latent variable. The use of I_n as dependent variables avoids the risk of endogeneity bias, while the use of a random component in α_n recognises our inability to accurately measure attitudes, perceptions and convictions. Finally, this model is directly applicable to forecasting, where, post estimation, the measurement component of the model can be removed, such that no role for the indicators I_n remains in the forecast calculations.

Experimental design, data and descriptive results

Table 1 show the various attributes used in the surveys, along with the levels for each attribute. Additionally, we show the expected direction in which a change in attribute levels would affect utility.

Data was collected at The Spine Centre of Southern Denmark, Lillebælt Hospital, Middelfart in the Region of Southern Denmark. This centre is the only public spine centre in the region, which has approximately 1.3 million inhabitants. At the centre, a range of HCPs, including physiotherapists, rheumatologists, psychologists and surgeons work together, treating approximately 12,500 new out-patients each year.

A total of 561 questionnaires were handed out to patients of which 348 were returned, corresponding to a response rate of 62 %. All patients receiving a questionnaire were included in a database and analysis of non-response amongst patients was performed using appropriate tests and showed no difference between the response and non-response groups in terms of mean values of age, gender and back- or leg pain. Patients giving missing information for any of the core variables included in the analysis were excluded, resulting in a final sample size of 297 respondents³.

The SC scenarios were developed using Ngene (Choicemetrics, 2010), using a Bayesian D-efficient design. Priors were obtained from a multinomial logit model based on quantitative pilot study with 17 responding patients each answering ten choice tasks, and through qualitative work including interviews with HCPs. For full details, see Kloejgaard, M.; Bech, M.; Soegaard, R. (2012). The final design contained 18 choice scenarios. To reduce respondent burden, these were split into three blocks of

³Core variables include the answers to the indicator questions, the choices, and the socio-demographic variables used. While missing socio-demographic variables could have been imputed, this can be unreliable, especially with a relatively modest sample size. For respondents with missing indicators and choices, we took the decision that it was preferable to remove such respondents completely rather than include them with only a limited set of dependent variables.

Table 1: Attributes and levels. First level for each attribute is baseline

Attribute	Levels	Hypothesis
Modality	Non-surgical	-/+
	Surgical	
Pain	Same	+ +
	Less	
	None	
Problems with activities of daily living (ADL)	Same	+ +
	Fewer	
	None	
Risk of relapse	1 in 10	- -
	2 in 10	
	3 in 10	
Time to treatment effect	1 month	- - -
	3 months	
	6 months	
	12 months	

six tasks, where orthogonal blocking was used to ensure no correlation between blocks and attribute combinations. The patients were randomly distributed between blocks and no significant differences concerning age, gender and mean pain values was observed between blocks. In each task, the survey presented respondents with three treatment options from which they were asked to indicate their preferred option, with the first two alternatives representing the hypothetical treatment options, and the remaining option being a no-choice option. An example choice set is shown in Figure 2.

MIRJA: HAVE CUT THE TEXT ABOVE THE SCREENSHOT

The attributes included in the survey reflected the treatment, its effects and risks as well as a time aspect, mirroring the differences in outcomes experienced by patients taking part in both surgery and non-surgical cross-disciplinary therapy (Bederman et al., 2010; Chou et al., 2009; Weiner and Essis, 2006). The qualitative work suggested that these attributes best reflected the complexity of the treatment choice faced by patients and also included the majority and most important aspects of the drivers for a choice. The levels used in the survey were based on qualitative and quantitative tests of different levels and were intended to ensure trade-offs while remaining realistic (Kloejgaard, M.; Bech, M.; Soegaard, R., 2012).

In addition to the answers to the SC questions, data on background, socio-economic characteristics and experience with and attitudes towards treatment options was gathered and is summarised in Table 2. The included questions were based on the validated Dallas Pain Questionnaire and on a widely used LBP scale. For the attitudinal indicators, we focused on current situation regarding pain and impact on their lives, as the literature points to more pain and everyday problems as a motivation for preferring surgery (Bederman et al., 2010; Bridwell et al., 2000; Lurie et al., 2008; Turner et al., 1998). A similar effect on preferences has been shown for patients with longer pathways and more experience with different non-surgical treatments (Lurie et al., 2008). For patient characteristics, we included a range of common socio-demographic questions as well as questions on recommendations of treatments, hypothesizing that patients were affected by the advice of others.

MIRJA: CAN YOU REDO THE GRAPH FOR FIGURE 3, EXCLUDING THE ANSWERS FOR THE HCPs, PLEASE?

The answers to attitudinal indicators are summarised in Figure 3. As can be seen, patients' answers to the six first indicator questions (I_{1-6}) regarding back- and leg pain show that current back pain is quite evenly spread on the scale, while, as expected, the worst possible pain experienced is quite

Table 2: Attitudinal questions and respondent characteristics. Variable name for indicators in brackets.

Attitudinal indicators	Scale
How much back pain do you feel right now? (I_1)	0 (none) – 10 (worst possible pain)
What is the worst back pain you have experienced in the past 2 weeks? (I_2)	0 (none) – 10 (worst possible pain)
What is the average back pain you have experienced in the past 2 weeks? (I_3)	0 (none) – 10 (worst possible pain)
How much leg pain do you feel right now? (I_4)	0 (none) – 10 (worst possible pain)
What is the worst leg pain you have experienced in the past 2 weeks? (I_5)	0 (none) – 10 (worst possible pain)
What is the average leg pain you have experienced in the past 2 weeks? (I_6)	0 (none) – 10 (worst possible pain)
How often do you use pain-killers? (I_7)	0 (never) – 10 (always)
How much does your pain negatively affect your sleep? (I_8)	0 (never) – 10 (always)
Have your back pain affected relationships with friends/family? (I_9)	0 (not at all) – 10 (dramatically)
How much physical support do you need from others? (I_{10})	0 (none) – 10 (all the time)
To what degree do you feel others are frustrated with your pain? (I_{11})	0 (not at all) – 10 (dramatically)
How many times have you visited the following? (I_{12})	Chiropractor, Physiotherapist, Rheumatologist, Other Specialist, Acupuncturer, Reflexologist, E.R., X-rays/Scans, Hospitalizations.
Do you have a preferred treatment? (I_{13})	Yes - surgery, Yes -non-surgical, No
Patient characteristics	Categories
Did your GP recommend a treatment?	Yes - surgery, Yes -non-surgical, No
Did another HCP recommend a treatment?	Yes - surgery, Yes -non-surgical, No
Did family/friends/fellow patients recommend a treatment?	Yes - surgery, Yes -non-surgical, No
Are you employed?	Yes, no
Have you had sick-leaves due to back pain?	Yes, no
What is your age?	18-44,45-54,55-64,64+
What is your yearly gross-income(DKR)?	<200.000, 200-400, >400.000

rare. Regarding average back pain, respondents score this to be less severe than the worst experienced pain, but more severe than their present state. The same picture is seen for leg pain, although in general leg pain is rated less severe. More than 20% state they use pain killers all the time (I_7), but a little less than 20% never use pain killers. The same pattern of heterogeneity is observed for sleep disturbances (I_8), where 10% feel they are always disturbed while a similar amount is never disrupted. Few respondents feel they need a lot of support from friends and family or that their conditions have influenced their relationships (I_{9-10}) and even fewer feel that they are met with frustration from their surroundings(I_{11}).

Results for the last two indicators are shown in Table 3. Most patients have visited a range of HCPs (I_{12}), especially physiotherapists and chiropractors, and most preferred non-surgical treatment when asked ex-ante(I_{13}). Table 3 further shows results of patients' characteristics. Patients were primarily low- to middle income and covered a broad age spectrum. Half the patients were employed and half

Table 3: Characteristics and attitudes. Categorical questions.

Variable	Results
Characteristics HCPs	
Years of experience	<5 years : 23 % 5-10 years: 31 % >10 years: 46 %
Private Employment	Yes: 3% No: 97 %
Type of HCP	Medical training: 26 % Others: 72 %
Indicators Patients	
How many times have you visited the following?(I_{12})	Chiropractor (Mean(SD)): 7.9 (6.2) Physiotherapist: 10.2 (16.6) Rheumatologist: 2.2 (4.3) Other Specialist: 1.8 (2.3) Acupuncturer: 4.1 (4.5) Reflexologist: 3.9 (5.1) E.R., X-rays/Scans, Hospital: 2.7 (2.7)
Do you have a preferred treatment? (I_{13})	Yes – surgery: 11 % Yes -non-surgical: 42 % No: 47 %
Characteristics Patients	
Did your GP recommend a treatment?	Yes – surgery: 4 % Yes -non-surgical: 32 % No: 64 %
Did another HCP recommend a treatment?	Yes – surgery: 3 % Yes -non-surgical: 20 % No: 77 %
Did family/friends/fellow patients recommend a treatment?	Yes – surgery: 6 % Yes -non-surgical: 28 % No: 66 %
Are you employed?	Yes: 49 % No: 51 %
Have you had sick-leaves due to back pain?	Yes: 51 % No: 49 %
What is your age?	18-44: 19 % 45-54: 28 % 55-64: 25 % 64+: 28 %
What is your yearly gross-income?	<200.000 DKR: 40 % 200.000-400.000 DKR: 47 % >400.000 DKR: 13 %

had taken sick-leave due to back pain. If patients had had any treatment recommendation from HCPs or their GP, it tended to be non-surgical. Patients' friends and families were more likely to have recommended treatment and also favoured non-surgical options.

Model specification

In this section, we explain how the model structure presented in the modelling methodology section was used in the present case study. We estimated four different models, namely a Multinomial Logit (MNL) model, a Mixed Multinomial Logit model (MMNL), and two hybrid model structures. We will now look at the specification of the four models in turn. A crucial component in the comparisons across models is a consistent treatment of socio-demographic characteristics, ensuring that the MNL and MMNL base structures equate to reduced form versions of the hybrid structure (cf. Vij and Walker, 2012). All models were coded and estimated using Ox 6.2 (Doornik, 2001). To try to avoid issues with local optima as much as possible, we ran the models with numerous different sets of starting values, in addition to each time using an approach that starts from the position of the best fitting set of starting values from a set of 1,000 randomly selected combinations. In the models with random coefficients, we made use of 500 MLHS draws per respondent and per random component (cf. Hess et al., 2006). Finally, the repeated choice nature of the data was recognised in the specification of the sandwich matrix for estimating robust standard errors (cf. Daly and Hess, 2011).

MNL model

In the MNL model, the following specification was initially used for the utility for alternatives 1 and 2 for both patients and HCPs.

$$\begin{aligned} V_{n,j,t} = & \delta_j \\ & + \beta_{\text{surgery}} \cdot x_{\text{surgery},n,j,t} \\ & + \Delta_{\text{GP recommended surgery}} \cdot z_{\text{GP recommended surgery},n} \\ & + \Delta_{\text{HCP recommended surgery}} \cdot z_{\text{HCP recommended surgery},n} \\ & + \Delta_{\text{friends or family recommended surgery}} \cdot z_{\text{friends or family recommended surgery},n} \\ & + \Delta_{\text{employed}} \cdot z_{\text{employed},n} \\ & + \Delta_{\text{previous sick leave}} \cdot z_{\text{previous sick leave},n} \\ & + \Delta_{\text{age 45-54}} \cdot z_{\text{age 45-54},n} \end{aligned}$$

$$\begin{aligned}
& + \Delta_{\text{income between DKK200,000 and DKK400,000}} \cdot z_{\text{income between DKK200,000 and DKK400,000},n} \\
& + \Delta_{\text{income above DKK400,000}} \cdot z_{\text{income above DKK400,000},n} \\
& + \beta_{\text{ADL fewer}} \cdot x_{\text{ADL fewer},n,j,t} \\
& + \beta_{\text{ADL none}} \cdot x_{\text{ADL none},n,j,t} \\
& + \beta_{\text{pain less}} \cdot x_{\text{pain less},n,j,t} \\
& + \beta_{\text{pain none}} \cdot x_{\text{pain none},n,j,t} \\
& + \beta_{\text{risk 20}} \cdot x_{\text{risk 20},n,j,t} \\
& + \beta_{\text{risk 30}} \cdot x_{\text{risk 30},n,j,t} \\
& + \beta_{\text{wait 3 months}} \cdot x_{\text{wait 3 months},n,j,t} \\
& + \beta_{\text{wait 6 months}} \cdot x_{\text{wait 6 months},n,j,t} \\
& + \beta_{\text{wait 12 months}} \cdot x_{\text{wait 12 months},n,j,t}
\end{aligned} \tag{8}$$

This specification applies to $j = 1, 2$, where δ_j is a constant estimated with $j = 1$ and fixed to zero (for identification) with $j = 2$. For the five attributes describing the alternatives, we dummy coded the attributes, where the base level was fixed to zero (for identification), which applies to non-surgical treatment⁴, same level of ADL, same level of pain, and the lowest level of risk or relapse (10 %) and waiting time (1 month). This thus initially led to the estimation of ten β parameters.

The decision to not use a continuous specification for weight time was motivated by a desire to investigate possible non-linearity in sensitivities⁵. This produced an interesting finding, with results showing no significant differences between the three estimated waiting time parameters, meaning that a single term is estimated in that group, namely $\beta_{\text{wait} > 1 \text{ month}}$, i.e. for waiting times of more than one month. It should be noted that this does not mean that patients ignore the waiting time attribute, given that all three initial estimates ($\beta_{\text{wait 3 months}}$, $\beta_{\text{wait 6 months}}$ and $\beta_{\text{wait 12 months}}$) were significantly different from zero. Rather, there was no significant difference between them, suggesting that patients have a threshold preference, treating all options with waiting times of over one month in the same

⁴It is not the case that each choice set includes exactly one surgery option and one non-surgery option, plus the opt out. In some of the tasks, there are two surgery options, while in others there are two non-surgery options. So treatment modality was an attribute rather than a label of the alternative alone.

⁵With only four levels for the attribute, any parameterization of the non-linearity would have suffered from the low number of support points on the marginal utility distribution. Clearly, estimating level specific coefficients gives the highest amount of information possible.

way.

We in addition estimated shifts in the preference for surgery for a number of key socio-demographic groups, testing the impact of past recommendations by either GPs, HCPs, or friends and family, the impact of employment status, past sick leave, being aged between 45 and 54, and falling in two different income groups. Other interactions were found not to be significant.

The utility function for the no treatment alternative was specified by a constant, such that:

$$V_{n,3,t} = \delta_{nc} \tag{9}$$

MMNL model

In the MMNL model, we allowed for random heterogeneity in the preference for surgery⁶. In particular, let $V_{n,j,t,\text{surgery}}$ be the part of the utility function that relates to the surgery attribute, i.e., in Equation 8, we would have that $V_{n,j,t,\text{surgery}} = \beta_{\text{surgery}} \cdot x_{\text{surgery},n,j,t}$. In the MMNL model, we replace this by:

$$\begin{aligned} V_{n,j,t,\text{surgery}} &= \beta_{\text{surgery}} \cdot x_{\text{surgery},n,j,t} \\ &+ \sigma_{\text{surgery}} \cdot \xi_{1,n} \cdot x_{\text{surgery},n,j,t} \end{aligned} \tag{10}$$

where all utility components other than those relating to the surgery attribute ($x_{\text{surgery},n,j,t}$) remain unchanged from the MNL utility function in Equation 8. With this specification, $\xi_{1,n}$ is a random variate that follows a standard Normal distribution across individual respondents but is held constant across choices for the same respondent. This ensures that the preference for surgery now follows a Normal distribution across respondents, with mean β_{surgery} and standard deviation σ_{surgery} . The choice of the Normal distribution was specifically motivated by the fact that some individuals may prefer surgery over non-surgical treatment, with the opposite applying to others.

Hybrid model

In the hybrid model, we hypothesized that patients have underlying perceptions regarding the choice of treatment modality and that these are drivers for the observed choices. Hence, we created a latent

⁶We acknowledge the potential existing of random heterogeneity also in other model parameters; the preference for surgery was simply the key parameter of interest, also in making the link to the later hybrid models.

variable focusing on treatment modality. The choice of latent variable was based on simple analysis showing that the modality attribute had a big impact on econometric models in both samples, a result in line with others' findings (Lurie et al., 2008). The presentation of the hybrid model structure focusses on three separate components, namely the specification of the structural equation for the latent variable, the specification of the measurement model, and the specification of the utility function in the choice model component.

Structural equation for latent variable

As outlined above, a single latent variable is used in our models, relating to an underlying pro-surgery attitude. For consistency with the choice model, we use the same socio-demographic characteristics in the latent variable, where the separate identification of two parameters associated with the same characteristic is ensured by the fact that for one of them, the value is driven by both the choice data and the indicator variables. In particular, we have that:

$$\begin{aligned}
\alpha_n = & \gamma_{\text{GP recommended surgery}} \cdot z_{\text{GP recommended surgery},n} \\
& + \gamma_{\text{HCP recommended surgery}} \cdot z_{\text{HCP recommended surgery},n} \\
& + \gamma_{\text{friends or family recommended surgery}} \cdot z_{\text{friends or family recommended surgery},n} \\
& + \gamma_{\text{employed}} \cdot z_{\text{employed},n} \\
& + \gamma_{\text{previous sick leave}} \cdot z_{\text{previous sick leave},n} \\
& + \gamma_{\text{age 45-54}} \cdot z_{\text{age 45-54},n} \\
& + \gamma_{\text{income between DKK200,000 and DKK400,000}} \cdot z_{\text{income between DKK200,000 and DKK400,000},n} \\
& + \gamma_{\text{income above DKK400,000}} \cdot z_{\text{income above DKK400,000},n} \\
& + \xi_{2,n},
\end{aligned} \tag{11}$$

where $\xi_{2,n}$ is defined as in Equation 2.

Specification of measurement model

A total of 13 indicators were used, as listed in Table 2. A decision was taken to use a continuous specification despite the categorical nature of the first 12 indicators. This is a potential limitation,

but could not be avoided as it would not have been realistic to use an ordered specification given the resulting proliferation of parameters with 10 levels for the indicators.

These 12 indicators were first centred on zero by subtracting the sample mean, after which their value could be explained using Equation 4, with two parameters estimated for indicator I_1 , namely ζ_{I_1} to measure the impact of the latent variable on the indicator, and σ_{I_1} for the standard deviation, where a constant is no longer needed following the centring on zero. For the first twelve indicators, we would thus have that:

$$P_{I_{n,l}} = \frac{1}{\sigma_{I,l}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{I_{n,l}-\zeta_l\alpha_n}{\sigma_l}\right)^2} \quad (12)$$

for $l = I_1, \dots, I_{12}$.

For the final indicator, we used a binary logit model to explain whether a patient has a preference for surgery, with a constant $\delta_{I_{13}}$ estimated alongside an impact of the latent variable, given by $\zeta_{I_{13}}$. This would thus give us:

$$P_{I_{n,13}} = \frac{(I_{n,13} == 1) e^{(\delta_{I_{13}} + \zeta_{I_{13}}\alpha_n)} + (I_{n,13}! = 1)}{1 + e^{(\delta_{I_{13}} + \zeta_{I_{13}}\alpha_n)}} \quad (13)$$

where $(I_{n,13} == 1)$ is equal to 1 if the respondent indicates a preference for surgery, and 0 otherwise, with the converse applying for $(I_{n,13}! = 1)$. The thirteen elements from Equation 12 and Equation 13 are then multiplied together to form the likelihood of the observed indicators in Equation 5. For HCPs, an equivalent version of Equation 12 is used, but with only three indicators.

Specification of utility in choice model

Two different version of the hybrid choice model were estimated, one as an extension of the MNL model, and one as an extension of the MMNL model. For the former, we now rewrite the component of utility relating to the surgery attribute as:

$$V_{n,j,t,\text{surgery}} = \beta_{\text{surgery}} \cdot x_{\text{surgery},n,j,t} + \tau_{\text{surgery}} \cdot x_{\text{surgery},n,j,t} \cdot \alpha_n, \quad (14)$$

while, for the MMNL model extension, we use:

$$\begin{aligned}
V_{n,j,t,\text{surgery}} &= \beta_{\text{surgery}} \cdot x_{\text{surgery},n,j,t} \\
&+ \sigma_{\text{surgery}} \cdot \xi_{1,n} x_{\text{surgery},n,j,t} \\
&+ \tau_{\text{surgery}} \cdot x_{\text{surgery},n,j,t} \cdot \alpha_n.
\end{aligned} \tag{15}$$

With either model, τ_{surgery} now measures a deviation from the mean preference for surgical treatment as a function of the latent attitude.

Results

The main estimation results are summarised in Table 4. Looking first at model fit, we can see that the MMNL model obtains an improvement of 85.09 units in log-likelihood over the MNL model, which, at the cost of just one additional parameter, is significant even at very high levels of confidence. The overall model fit for the two hybrid models cannot be directly compared to that for the MNL and MMNL models, as it relates to both the choice data and the explanation of the indicator variables. We observe an improvement in log-likelihood by 75.4 units when moving from the first hybrid choice model to the second, where this is highly significant at the cost of just one additional parameter (σ_{surgery}). The improvement is a little smaller than what we see when moving from MNL to MMNL, a reflection of the fact that some of the heterogeneity across respondents in the preference for surgery is now also explained by the latent variable α .

It is also possible to factor out the portion of the log-likelihood that relates only to the choice data, i.e. the first component in Equation 6, relating to P_{C_n} . From this, we can observe that the explanation of the choice data is better in Hybrid₁ than in MNL, while it is equivalent in Hybrid₂ and MMNL, bar simulation noise. We also see that the explanation of the choices in Hybrid₁ is below that of MMNL. All three findings are entirely consistent with intuition, as we will now explain. Firstly, the better log-likelihood for the choice component in Hybrid₁ compared to the MNL model is to be expected given the additional parameter τ , which now allows for random heterogeneity. Secondly, the lower log-likelihood for the choice component in Hybrid₁ compared to the MMNL model is to be expected given that, in the latter, the random heterogeneity in the preference to surgery is freely estimated on

the choice data alone, while, in the former, it is constrained by the joint maximisation on the choice data and the indicators. Finally, MMNL and Hybrid₂ would indeed be expected to offer the same log-likelihood on the choice data, given that, with either, the heterogeneity in the preference for surgery follows a Normal distribution (a sum of two Normals in Hybrid₂), and with the same flexibility being incorporated in relation to deterministic heterogeneity. This is in line with theory, as discussed by Vij and Walker (2012). This then raises the question as to the actual benefit of the hybrid framework, and this is twofold. Firstly, there is a gain in efficiency by making use of additional data. Secondly, we are able to decompose the heterogeneity into a part related to attitudes as captured by the attitudinal questions, and a remaining purely random part. We will return to both points below.

We first focus on the main effects across all the models. We note a negative estimate for δ_1 , while the estimate for δ_{nc} is not significant in any of the models. This indicates that, all else being equal, and with the middle alternative being the base ($\delta_2 = 0$), we observe a slight preference for choosing the second alternative. The parameter associated with surgery is negative and significant across all models, indicating a preference for non-surgical treatment over surgery. We also observe that respondents have positive utilities for less pain and even more so for no pain, as well as for fewer and even more so for no problems with activities of daily living. This preference is stronger for pain reductions than for problems with activities of daily living. An increase in risk of relapse from 10% to 20% is not statistically significant in any of the models, while an increase to 30% is highly significant and valued negatively in all models. This justifies avoiding a linear specification for this attribute. As already mentioned earlier, we see a constant sensitivity against waiting times greater than 1 month.

We next turn to the heterogeneity in the preference for surgery, focussing first on the parameters relating only to the choice model component. We first see positive impacts of others recommending surgery, where these are however only significant at usual levels for GPs in the MMNL model, and for friends and family in both the MNL and MMNL models. The impact of GPs' recommendations is most positive, followed by friends and family and then HCPs. None of the other socio-demographic effects (employment status, past sick leave, age and income) shows significant effects in any of the models. Finally, the estimate for σ_{surgery} is highly significant in both the MMNL and Hybrid₂ models, showing the presence of random variations in the preference for surgery. With the estimate for β_{surgery} being almost identical in the two models, we can see that the degree of relative heterogeneity is lower in the Hybrid₂ model; this is a direct reflection of the fact that some of that heterogeneity is now

Table 4: Estimation results

	MNL		MMNL		Hybrid ₁		Hybrid ₂	
Overall log-likelihood	-1,502.23		-1,417.14		-10,886.30		-10,810.90	
Choice component log-likelihood	-1,502.23		-1,417.14		-1,480.69		-1,417.23	
model parameters	18		19		53		54	
δ_1	-0.1236	-2.18	-0.1813	-2.65	-0.1303	-2.26	-0.1819	-2.67
δ_{nc}	0.0081	0.05	0.0653	0.34	0.0005	0.00	0.0623	0.33
$\beta_{surgery}$	-1.1327	-5.25	-1.7476	-5.77	-1.1813	-5.56	-1.7485	-5.81
$\beta_{ADL\text{ fewer}}$	0.5837	5.25	0.6605	5.00	0.5915	5.30	0.6611	5.02
$\beta_{ADL\text{ none}}$	0.8516	9.03	1.0167	8.81	0.8537	8.95	1.0117	8.80
$\beta_{pain\text{ less}}$	1.3742	13.34	1.5306	12.81	1.3880	13.25	1.5277	12.83
$\beta_{pain\text{ none}}$	1.6324	14.91	1.9693	15.00	1.6653	14.89	1.9699	15.02
$\beta_{risk\ 20}$	-0.0320	-0.32	0.0250	0.21	-0.0201	-0.20	0.0334	0.28
$\beta_{risk\ 30}$	-0.3106	-3.27	-0.2717	-2.45	-0.3163	-3.32	-0.2709	-2.46
$\beta_{wait > 1\text{ month}}$	-0.6356	-8.00	-0.7516	-8.25	-0.6551	-8.17	-0.7556	-8.28
$\Delta_{GP\text{ recommended surgery}}$	1.0012	1.87	1.5078	2.03	0.7925	1.52	1.3533	1.79
$\Delta_{HCP\text{ recommended surgery}}$	0.5046	0.60	0.6230	0.51	0.2213	0.28	0.3417	0.28
$\Delta_{\text{friends or family recommended surgery}}$	0.8040	2.37	1.2172	2.50	0.5304	1.48	0.8854	1.75
Δ_{employed}	-0.1617	-0.73	-0.2805	-0.94	0.0173	0.08	-0.0728	-0.25
$\Delta_{\text{previous sick leave}}$	0.1541	0.78	0.2021	0.73	-0.0357	-0.18	-0.0238	-0.08
$\Delta_{\text{age 45-54}}$	0.2187	1.02	0.3227	1.08	0.1014	0.46	0.2099	0.70
$\Delta_{\text{income between DKK200,000 and DKK400,000}}$	-0.0721	-0.29	0.0177	0.05	0.0200	0.08	0.1079	0.33
$\Delta_{\text{income above DKK400,000}}$	0.0970	0.27	0.2130	0.44	0.2399	0.67	0.3750	0.80
$\sigma_{surgery}$			1.7726	12.18			1.7102	12.12
$\gamma_{GP\text{ recommended surgery}}$					0.3237	0.91	0.3307	1.03
$\gamma_{HCP\text{ recommended surgery}}$					0.5197	1.62	0.5153	1.67
$\gamma_{\text{friends or family recommended surgery}}$					0.7746	3.41	0.7434	3.38
γ_{employed}					-0.4725	-3.01	-0.4962	-3.15
$\gamma_{\text{previous sick leave}}$					0.4889	4.08	0.4747	3.86
$\gamma_{\text{age 45-54}}$					0.3119	2.10	0.3390	2.19
$\gamma_{\text{income between DKK200,000 and DKK400,000}}$					-0.2609	-1.68	-0.2308	-1.48
$\gamma_{\text{income above DKK400,000}}$					-0.3656	-1.77	-0.3390	-1.64
$\tau_{surgery}$					0.4043	3.10	0.4456	2.81
ζ_{I_1}					1.7322	12.16	1.7361	10.97
ζ_{I_2}					1.4804	8.54	1.4774	7.95
ζ_{I_3}					1.6286	11.12	1.6337	10.13
ζ_{I_4}					1.5446	7.45	1.5866	7.25
ζ_{I_5}					1.5106	5.55	1.5584	5.39
ζ_{I_6}					1.5166	7.31	1.5576	7.08
ζ_{I_7}					1.9803	11.72	2.0033	11.71
ζ_{I_8}					1.7183	11.22	1.7419	11.20
ζ_{I_9}					1.9625	14.40	1.9974	14.46
$\zeta_{I_{10}}$					1.7847	12.97	1.8162	13.00
$\zeta_{I_{11}}$					1.5450	10.32	1.5753	10.44
$\zeta_{I_{12}}$					0.0980	1.19	0.0942	1.14
$\zeta_{I_{13}}$					0.4267	5.21	0.4098	5.15
σ_{I_1}					1.3557	8.72	1.3769	8.15
σ_{I_2}					1.7235	12.74	1.7432	11.97
σ_{I_3}					1.2648	7.73	1.2841	7.18
σ_{I_4}					2.1088	12.62	2.0857	11.68
σ_{I_5}					2.7349	15.46	2.7133	14.27
σ_{I_6}					2.1162	13.22	2.0944	12.16
σ_{I_7}					3.0760	25.43	3.0738	25.32
σ_{I_8}					2.5681	21.75	2.5629	21.76
σ_{I_9}					2.3156	17.36	2.2996	17.09
$\sigma_{I_{10}}$					2.2687	19.56	2.2564	19.48
$\sigma_{I_{11}}$					2.3327	20.05	2.3211	19.92
$\sigma_{I_{12}}$					1.3145	11.98	1.3149	11.99
$\delta_{I_{13}}$					-2.2784	-11.30	-2.2659	-11.32

captured by the latent variable.

The benefits of the hybrid structures in terms of efficiency are clear to see when looking at the estimates and significance levels for the socio-demographic terms in the latent variable (γ). We now observe significant positive effects on the latent variable for recommendations from friends and family, where, contrary to the estimates from the Δ terms in the choice model component, we note that these recommendations matter more than those from HCPs and GPs, with the lowest value attached to the latter. Furthermore, we observe a negative and significant impact on the attitude towards surgery for respondents in employment, presumably due to the implied disruption to work life. This is also consistent with the positive estimate for respondents who have had to take sick leave in the past, where this is again significant in both models. We see a significant positive impact on the latent attitude for respondents in the middle age group, while, albeit not significant at usual levels, the attitude towards surgery becomes more negative with higher income. In both of the hybrid structures, the estimate for τ_{surgery} is positive and significant, showing that a more positive latent attitude leads to a greater preference for surgery in the choice model component.

As discussed earlier, the latent variable also includes a random disturbance, which follows a standard Normal distribution. This then means that the latent variable also contributes to the random heterogeneity in β_{surgery} in the choice model. Some important observations can be made here. Remembering that the coefficient of variation ($\frac{\sigma_{\text{surgery}}}{\beta_{\text{surgery}}}$) in the MMNL model was equal to 1.01, we can see a drop in the heterogeneity in the Hybrid₁ model to 0.34. This is again a direct reflection of the fact that in the latter, the estimation of the heterogeneity in the choice model component is constrained by the fact that any heterogeneity needs to be perfectly correlated with the heterogeneity in the measurement model. In the Hybrid₂ model, this requirement disappears, and, as a result, we see essentially the same level of overall heterogeneity in the choice model component as in the MMNL model, with $\frac{\sqrt{\sigma_{\text{surgery}}^2 + \tau_{\text{surgery}}^2}}{\beta_{\text{surgery}}} = 1.01$, in line with theory. It also becomes clear that only a small share of the overall heterogeneity is in this case linked to the latent variable, which is responsible for just over 6% of the total variance. Two possible interpretations arise. Firstly, there is a possibility that the heterogeneity in the preferences towards surgery is largely unrelated to attitudes and perceptions, and this would line up nicely with the reasoning that there exists substantial uncertainty as to the benefits of either treatment modality. Secondly, there is a possibility that the attitudinal statements used in our work capture attitudes that are not directly linked to heterogeneity in preferences for surgery. In reality, a

mixture of the two potentially applies, though we wish to highlight again the background work that went into formulating the attitudinal questions.

The final component of the hybrid structure, the measurement model, explains the impact of the latent variable on the range of indicator variables, where the labelling in Table 4 is that from Table 2. The 6 indicators (I_{1-6}) of back- and leg pain all have a positive and significant associated ζ parameter, meaning that a more positive latent pro-surgery attitude is associated with higher stated pain levels. The same pattern is seen for intake of pain-killers (I_7) and disturbances during sleep (I_8). Higher pro-surgery latent variables are also associated with a greater perceived impact of LBP on relationships with family and friends (I_9), the perceived need for support (I_{10}) and the perceived level of frustration caused to others (I_{11}). The number of visits to HCPs (I_{12}) is only weakly linked to the latent attitude, but a higher latent pro-surgery attitude is strongly linked to a higher probability of stating a preference for surgery (I_{13}), where the negative associated $\delta_{I_{13}}$ term indicates an overall preference for non-surgical treatment (cf. Equation 13).

Substitution Rates

As a final step in our analysis, we now proceed with a further comparison of results across the four models. To avoid issues with scale differences, this comparison is carried out with the help of marginal rates of substitution, i.e. looking at the change in an attribute required to compensate for a change in another attribute to keep the total utility constant. Traditionally, researchers calculate marginal willingness to pay, using respondents' marginal utility of price as a denominator in ratios of coefficients. In health economics, time is often used as a payment-vehicle instead (Gerard et al., 2004; Gerard and Lattimer, 2005; Ratchliffe, 2000; Ryan et al., 2001; Yi et al., 2011), as a price attribute would not mimic the real world scenario of no direct user payment for the provision of health care in many European settings.

As a MRS is basically just a ratio between the marginal utility of two attributes, examining the relative importance of attributes to one another, any coefficient can in principle be used as the denominator. The use of a time coefficient is generally motivated by the continuous linear treatment of the associated attribute and researchers rarely test whether this assumption of linearity holds, and thus potentially calculate MRS based on incorrect premises. In this data, the sensitivity to the time attribute was found not to be linear, as a simple step function emerged. As a result, the use of time as the base

Table 5: MRS
(a) MRS related to surgery

	MNL	MMNL	Hybrid ₁	Hybrid ₂
25 th percentile	3.15	2.42	2.43	2.50
median	3.65	6.89	3.49	6.94
mean	3.21	6.91	3.34	6.95
std. dev.	1.24	6.63	1.75	6.58
75 th percentile	3.88	11.38	4.47	11.39
mean for patients with recommendation by GP, practitioner and friends & family, not employed, with past sick leave, aged 45-54 and lowest income group	-4.99	-7.82	-4.45	-7.71
mean for patients with recommendation only by practitioner, employed, no past sick leave, aged 45-54 and highest income group	2.23	4.39	2.63	4.60
mean for patients without recommendation, employed, no past sick leave, highest income group	3.85	6.68	3.99	6.71

(b) Other MRS

	MNL	MMNL	Hybrid ₁	Hybrid ₂
fewer ADL problems	-1.88	-2.43	-1.87	-2.44
no ADL problems	-2.74	-3.74	-2.70	-3.73
less pain	-4.42	-5.63	-4.39	-5.64
no pain	-5.26	-7.25	-5.26	-7.27
wait > 1 month	2.05	2.77	2.07	2.79

of comparison was not appropriate. Instead, we used the risk of relapse of 30% as the denominator, meaning that our MRS measures give the relative importance of a change in a given attribute compared to an increase from 10% to 30% in the risk of relapse. This limits us from comparing MRS to other studies, which given the specificity of the case and the limited literature is not easy in any case.

The results of these calculations are summarised in Table 5. For all attributes, the MRS was simply calculated as the ratio between the associated coefficient and the coefficient for a risk of 30%. Given the normalisations used for the different attributes, these MRS thus present the relative sensitivity to a move away from the current level of ADL, the current level of pain, and a waiting time of 1 month respectively, compared to an increase in the risk of relapse from 10% to 30%. For surgery, it presents the relative sensitivity of moving from non-surgical to surgical treatment, again relative to that increase in the risk of relapse. For this specific MRS, we obtain a distribution across respondents, as a function of both the socio-demographic interactions and the random components, each time in both the choice model component and the latent variable. Table 5(a) presents details for the resulting distribution, as well as mean values for three representative types of patients.

The coefficient associated with an increase in risk to 30% is obviously negative, and as a result, positive MRS values are obtained for attribute levels that are similarly associated with a loss in utility

(e.g. increased waiting time, or surgery vs non-surgery) while negative MRS values are obtained for desirable attribute levels (i.e. more desirable ADL and pain levels). What is more important than the sign is whether the actual value is greater or smaller than unity in absolute terms. A $|MRS| > 1$ implies that the change in utility resulting from moving away from the baseline of that given attribute is greater than that of the increase in risk from 10% to 30%. The opposite applies for $|MRS| < 1$.

Looking at the MRS derived from the MNL estimates, a move from non-surgical to surgical treatment is on average valued around three times as negatively as a risk increase from 10% to 30%. For the three types of representative individuals (out of the 256 different combinations of socio-demographic groupings), we first look at the type of patient with the highest pro-surgery attitude, namely one who has been recommended surgery from both HCPs and family, is unemployed and middle-aged, belongs to the lowest income group and has experience with sick leaves. This type of patient actually has a strong preference of surgical over non-surgical treatment, all else being equal, where the absolute preference is five times as strong as the sensitivity to a move from the lowest to the highest risk level. At the other extreme, we have a type of patient so averse to surgery that a move from non-surgical to surgical treatment is valued almost five times as negatively as the increase in risk from 10% to 30%, this being a patient who has not had any recommendations of surgery, is employed and is in the highest income group. Finally, the MRS is smaller in absolute value for a patient with a surgery recommendation only from a practitioner and who is employed and in the middle age group as well as highest income group, but who has not had any past sick leave. Improvements in ADL or pain are valued more positively than increases in risk are valued negatively, where these differences are more substantial for pain. An increase in waiting time beyond 1 month is worse than an increase in risk from 10% to 30% by a factor of almost two.

Turning to the MMNL model, we see a much broader range for the MRS related to surgery across respondents, an obvious result of incorporating random heterogeneity in preferences. Additionally however, we observe differences in those MRS measures not related to surgery, an indication that failing to capture the heterogeneity in the sensitivity to surgery can also impact on other parameters. As would be expected from theory and given the main estimation results in Table 4, there is essentially no difference between the MRS for the MMNL and Hybrid₂ models. Similarly, the Hybrid₁ MRS not related to surgery are very similar to the MNL ones, while we see a slightly broader range for the surgery MRS in the Hybrid₁ model as a result of incorporating the random component in the latent

variable. The extent of variation is smaller than in the MMNL and Hybrid₂ models, given the smaller impact by τ_{surgery} compared to σ_{surgery} .

Discussion and Conclusion

OBVIOUSLY, THEY ARE THE SAME ALSO IN MRS, BUT IN PRACTICE, ANALYSTS MAY DROP INSIGNIFICANT SOCIOS!

In researching decision making in health care amongst both HCPs and patients, it is recognized that attitudes, beliefs and perceptions have a substantial impact on the choices made of decisions taken. Researchers often collect data concerning these factors and sometimes include answers to such subjective questions in choice models. Crucially, the simple treatment of such response as explanatory variables inside a choice model, as is typically done in health economics, can expose an analyst to problems with endogeneity bias and measurement error.

This paper has offered detailed information on how a joint model, integrating observations of choices and answers to subjective questions, is designed. By using SC data collected in the difficult field of treatment of LBP, we show that patients' choices of treatment for their condition, and doctors' preferences for different treatment options and hence the advice they give, can be partly described by underlying latent attitudes. Our results suggest that not accounting for underlying attitudes might produce less accurate results. Indeed, the hybrid model used in this paper provides further insights into preferences and explains the drivers of attitudes that influence these preferences.

The empirical results are interesting and valid on their own, but are not surprising. A link between higher levels of pain and more negative effects on life and a pro-surgery attitude makes intuitive sense. The same can be said about a link between being employed and having higher incomes and a more negative attitude towards surgery. Similarly, the result that medically trained HCPs are more in favour provides a sensible explanation. Interestingly though, results point to a negative effect on the pro-surgery attitude for more experienced HCPs. This could indicate that HCPs work to learn or experience little or no effect of surgery. Equally captivating is the results of no effects of surgical recommendations from professionals to patients while recommendations by friends and family have a significant impact. This suggests that patients are significantly influenced by peers, a finding which might not be surprising but which is often overlooked or not measured in the literature on preferences

in health economics. In this particular case, with unclear evidence in terms of outcomes of any particular treatment, this finding is less worrisome than what it could be in other settings.

Future research should be dedicated to looking into the influence of peers on patients' preferences and joint decision making of patients and their families or friends. Any influence should also be accounted for in our models. It seems plausible to suggest that the effect found in this study is not context specific.

Interestingly, our results also show that HCPs are quite successful in knowing their patients' preferences as mean MRS estimates are similar for both HCPs and patients, albeit with an overestimation of the sensitivity to risk.

All in all, results show a strong impact on choices and substitution patterns by the latent factor. It seems clear that in this setting, HCPs have a very difficult task in guiding their patients and communication on expectations, attitudes and possibilities from both HCPs and patients is key. The study also provides additional information to the policy discussion about treatment choices and surgery rates and suggests that patients' choices of treatment modality is multifaceted.

INCORPORATE: Response: From our review of the literature and based on initial results for both patients and HCPs, treatment modality was clearly a very important issue the main issue driving choices. Influences of attitudes on other factors are likely to exist, but the data collected was mainly focused on attitudes towards/experience with choosing between treatment options. As a result, Hence the choice of latent variable was data-driven as well as based on a thorough literature review and expert interviews etc.. We do however now also acknowledge this as a potential area for future work. We believe that the type of model used in this paper shows great promise for future studies in health economics. Integrating choices with latent variables is arguably of even greater importance in cases of confusing evidence, such as LBP, but it seems evident that beliefs are utterly important in decision making in health care in general. It is equally evident that hybrid models have substantial advantages in terms of explaining heterogeneity opening the way for more precise willingness to pay measures or substitution rates, and ultimately better policy recommendations. As with the choice of attributes and levels, researchers should carefully choose indicators and characteristics to include in the hybrid models as part of thorough qualitative work, prior to designing choice surveys. Future studies should look at different health topics, include more data from HCPs and also look at the use of multiple latent attitudes in a single model.

Acknowledgements

The authors would like to thank staff and patients at Middelfart Spine Centre with a special thanks to Professor Claus Manniche. The authors would also like to thank Assistant Professor Line Bjørnskov Pedersen for valuable comments.

The first author also gives thank the Danish Research Council and CeSpine (<http://www.cespine.dk/>) for financial support.

References

- Abou-Zeid, M., Ben-Akiva, M., Bierlaire, M., Choudhury, C., Hess, S., 2010. Attitudes and Value of Time Heterogeneity. In: Van de Voorde, E., Vanellander, T. (Eds.), *Applied Transport Economics - A Management and Policy Perspective*. De Boeck Publishing, forthcoming, pp. 525–548.
- Allen, T., Rihn, J., Glassman, S., Currier, B., Albert, T., Phillips, F., 2009. An Evidence-Based Approach to Spine Surgery. *American Journal of Medical Quality* 24, 15–24.
- Balagué, F., Mannion, A. F., Pellisé, F., Cedraschi, C., Mar. 2012. Non-specific low back pain. *Lancet* 379 (9814), 482–91.
- Bederman, S., Mahomed, N. N., Kreder, H. J., McIsaac, W. J., Coyte, P. C., Wright, James, G., Jan. 2010. In the eye of the beholder: preferences of patients, family physicians, and surgeons for lumbar spinal surgery. *Spine* 35 (1), 108–15.
- Ben-Akiva, M., McFadden, D., Garling, T., Gopinath, D., Walker, J. L., Bolduc, D., Boersch-supan, A., Delquié, P., Larichev, O., Morikawa, T., Polydoropoulou, A., Rao, V., 1999. Extended Framework for Modeling Choice Behavior. *Marketing Letters* 10 (3), 187–203.
- Ben-Akiva, M., Walker, J., Bernardino, A., Gopinath, D., Morikawa, T., Polydoropoulou, A., 2002a. Integration of choice and latent variable models. In: Mahmassani, H. (Ed.), *In Perpetual motion: Travel behaviour research opportunities and application challenges*. Pergamon, Ch. 13, pp. 431–470.
- Ben-Akiva, M., Walker, J., McFadden, D., Train, K., Bhat, C. R., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D. S., Daly, A., de Palma, A., Gopinath, D., Karlstrom, A., Munizaga, M. A., 2002b. Hybrid choice models: Progress and challenges. *Marketing Letters* 13 (3), 163–175.
- Bishop, A., Foster, N. E., Thomas, E., Hay, E. M., Mar. 2008. How does the self-reported clinical management of patients with low back pain relate to the attitudes and beliefs of health care practitioners? A survey of UK general practitioners and physiotherapists. *Pain* 135 (1-2), 187–95.
- Bolduc, D., Alvarez-Daziano, R., 2010. On estimation of hybrid choice models. In: Hess, S., Daly, A. (Eds.), *Choice Modelling: The State-of-the-Art and the State-of-Practice: Proceedings from the Inaugural International Choice Modelling Conference*. Emerald Group Publishing, pp. 259–288.

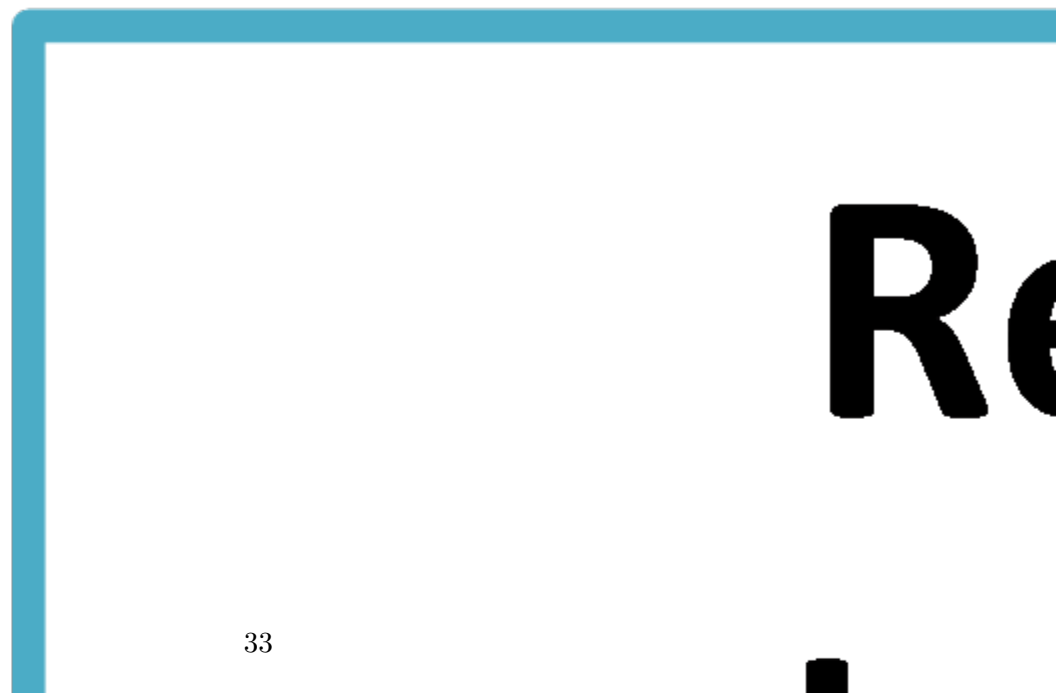
- Bolduc, D., Ben-Akiva, M., Walker, J., Michaud, A., 2005. Hybrid choice models with logit kernel: Applicability to large scale models. Emerald Group Publishing, pp. 275–302.
- Bridwell, K. H., Shufflebarger, H. L., Lenke, L. G., Lowe, T. G., Betz, R. R., Bassett, G. S., Sep. 2000. Parents’ and Patients’ Preferences and Concerns in Idiopathic Adolescent Scoliosis. *Spine* 25 (18), 2392–2399.
- Chocometrics, 2010. NGene 1.1. <http://www.choice-metrics.com/>.
- Chou, R., Baisden, J., Carragee, E. J., Resnick, D. K., Shaffer, W. O., Loeser, J. D., May 2009. Surgery for low back pain: a review of the evidence for an American Pain Society Clinical Practice Guideline. *Spine* 34 (10), 1094–109.
- Corbett, M., Foster, N., Ong, B. N., Oct. 2009. GP attitudes and self-reported behaviour in primary care consultations for low back pain. *Family practice* 26 (5), 359–64.
- Coudeyre, E., Rannou, F., Tubach, F., Baron, G., Coriat, F., Brin, S., Revel, M., Poiraudou, S., Oct. 2006. General practitioners’ fear-avoidance beliefs influence their management of patients with low back pain. *Pain* 124 (3), 330–7.
- Daly, A., Hess, S., 2011. Simple methods for panel data analysis. paper presented at the 90th Annual Meeting of the Transportation Research Board, Washington, D.C.
- Daly, A., Hess, S., Patrini, B., Potoglou, D., Rohr, C., Jul. 2012. Using ordered attitudinal indicators in a latent variable choice model: a study of the impact of security on rail travel behaviour. *Transportation* 39 (2), 267–297.
- Daziano, R. A., Bolduc, D., Mar. 2011. Incorporating pro-environmental preferences towards green automobile technologies through a Bayesian hybrid choice model. *Transportmetrica* 9 (1), 1–33.
- de Bekker-Grob, E. W., Ryan, M., Gerard, K., Feb. 2012. Discrete choice experiments in health economics: a review of the literature. *Health economics* 21 (2), 145–72.
- Doornik, J. A., 2001. *Ox: An Object-Oriented Matrix Language*. Timberlake Consultants Press, London.

- Gerard, K., Lattimer, V., Feb. 2005. Preferences of patients for emergency services available during usual GP surgery hours: a discrete choice experiment. *Family practice* 22 (1), 28–36.
- Gerard, K., Lattimer, V., Turnbull, J., Smith, H., George, S., Brailsford, S., Maslin-Prothero, S., Nov. 2004. Reviewing emergency care systems 2: measuring patient preferences using a discrete choice experiment. *Emergency medicine journal : EMJ* 21 (6), 692–7.
- Gibson, J N A. & Waddell, G., Jan. 2005. Surgery for degenerative lumbar spondylosis.
- Hess, S., Beharry-Borg, N., 2012. Accounting for latent attitudes in willingness-to-pay studies : the case of coastal water quality improvements in Tobago. *Environmental and Resource Economics* 52, 109–131.
- Hess, S., Train, K., Polak, J. W., 2006. On the use of a Modified Latin Hypercube Sampling (MLHS) method in the estimation of a Mixed Logit model for vehicle choice. *Transportation Research Part B* 40 (2), 147–163.
- Hole, A. R., Jul. 2008. Modelling heterogeneity in patients’ preferences for the attributes of a general practitioner appointment. *Journal of health economics* 27 (4), 1078–94.
- Houben, R. M. A., Ostelo, R. W. J. G., Vlaeyen, J. W. S., Wolters, P. M. J. C., Peters, M., Stomp-van den Berg, S. G. M., Apr. 2005. Health care providers’ orientations towards common low back pain predict perceived harmfulness of physical activities and recommendations regarding return to normal activity. *European journal of pain (London, England)* 9 (2), 173–83.
- Houben, R. M. A., Vlaeyen, J. W. S., Peters, M., Ostelo, R. W. J. G., Wolters, P. M. J. C., Stomp-van den Berg, S. G. M., 2004. Health care providers’ attitudes and beliefs towards common low back pain: factor structure and psychometric properties of the HC-PAIRS. *The Clinical journal of pain* 20 (1), 37–44.
- Irwin, Z. N., Hilibrand, A., Gustavel, M., McLain, R., Shaffer, W., Myers, M., Glaser, J., Hart, R. A., Oct. 2005a. Variation in surgical decision making for degenerative spinal disorders. Part I: lumbar spine. *Spine* 30 (19), 2208–13.
- Irwin, Z. N., Hilibrand, A., Gustavel, M., McLain, R., Shaffer, W., Myers, M., Glaser, J., Hart, R. A.,

- Oct. 2005b. Variation in surgical decision making for degenerative spinal disorders. Part II: cervical spine. *Spine* 30 (19), 2214–9.
- Kloejgaard, M.; Bech, M.; Soegaard, R., 2012. Designing a Stated Choice Experiment : The Value of a Qualitative Process. *Journal of Choice Modelling* 5 (2), 1–18.
- Linton, S. J., Vlaeyen, J., Ostelo, R., Dec. 2002. The back pain beliefs of health care providers: are we fear-avoidant? *Journal of occupational rehabilitation* 12 (4), 223–32.
- Lurie, J. D., Berven, S. H., Gibson-chambers, J., Tosteson, T., Tosteson, A., Hu, S. S., Weinstein, J. N., 2008. Patient Preferences and Expectations for Care Determinants in Patients With Lumbar Intervertebral Disc Herniation. *Spine* 33 (24), 2663–2668.
- Main, C. J., Foster, N., Buchbinder, R., May 2010. How important are back pain beliefs and expectations for satisfactory recovery from back pain? Best practice & research. *Clinical rheumatology* 24 (2), 205–17.
- Pincus, T., Foster, N. E., Vogel, S., Santos, R., Breen, A., Underwood, M., May 2007. Attitudes to back pain amongst musculoskeletal practitioners: a comparison of professional groups and practice settings using the ABS-mp. *Manual therapy* 12 (2), 167–75.
- Poiraudeau, S., Rannou, F., Le Henanff, A., Coudeyre, E., Rozenberg, S., Huas, D., Martineau, C., Jolivet-Landreau, I., Revel, M., Ravaud, P., Jun. 2006. Outcome of subacute low back pain: influence of patients' and rheumatologists' characteristics. *Rheumatology (Oxford, England)* 45 (6), 718–23.
- Ratcliffe, J., Jan. 2000. The use of conjoint analysis to elicit willingness-to-pay values. Proceed with caution? *International journal of technology assessment in health care* 16 (1), 270–5.
- Ryan, M., Bate, a., Eastmond, C. J., Ludbrook, a., Oct. 2001. Use of discrete choice experiments to elicit preferences. *Quality in health care : QHC* 10 Suppl 1 (Suppl I), i55–60.
- Turner, J. A., LeResche, L., Von Korff, M., Ehrlich, K., Feb. 1998. Back pain in primary care. Patient characteristics, content of initial visit, and short-term outcomes. *Spine* 23 (4), 463–9.
- van Tulder, M., Koes, B., Bombardier, C., Dec. 2002. Low back pain. Best practice & research. *Clinical rheumatology* 16 (5), 761–75.

- Vij, A., Walker, J., 2012. Hybrid choice models: Holy grail... or not? paper presented at the 13th International Conference on Travel Behaviour Research, Toronto.
- Vij, A., Walker, J., 2013. Handbook of Choice Modelling. Edward Elgar, Ch. Hybrid choice models: The identification problem.
- Weiner, B. K., Essis, F. M., Dec. 2006. Patient preferences regarding spine surgical decision making. Spine 31 (24), 2857–60; discussion 2861–2.
- Yi, D., Ryan, M., Campbell, S., Elliott, A., Torrance, N., Chambers, A., Johnston, M., Hannaford, P., Smith, B. H., May 2011. Using discrete choice experiments to inform randomised controlled trials: an application to chronic low back pain management in primary care. European journal of pain (London, England) 15 (5), 531.e1–10.

Figure 1: Structure of a hybrid model approach



You are now asked to make a series of hypothetical treatment choices. In each task you can either choose treatment option A or B or none.

Treatment A	Treatment B
The treatment is surgery	The treatment is cross-disciplinary therapy
After treatment your pain will be unchanged	After treatment, you'll have no pain
After treatment you'll have fewer problems with activities of daily living	After treatment you'll have the same problems with activities of daily living
The risk of relapse is 1 in 10	The risk of relapse is 3 in 10
It will take 3 months for the treatment to work	It will take 12 months for the treatment to work

I prefer (choose one) Treatment A _____ or Treatment B _____

I do not want any of the treatments _____

Figure 2: Choice set as shown to patients

Figure 3: Attitudes. 0-10 scaled questions as described in Table 2

