HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY of York

WP 14/12

Modelling nonlinearities and reference-dependence in general practitioners' income preferences

Jon H. Holte; Peter Sivey; Birgit Abelsen & Jan Abel Olsen

August 2014

york.ac.uk/res/herc/hedgwp

Modelling nonlinearities and reference-dependence in general practitioners' income preferences

Jon H. Holte[1]*, Peter Sivey [2], Birgit Abelsen[1], Jan Abel Olsen[1]

[1] Department of Community Medicine, University of Tromsø

[2] School of Economics, La Trobe University, Melbourne, Australia

Abstract

This paper tests for the existence of nonlinearity and reference-dependence in income preferences for general practitioners. Confirming the theory of reference dependent utility, within the context of a discrete-choice experiment, we find that *losses loom larger than gains* in income for Norwegian GPs, a 10% decrease in income is valued approximately equal to a 30% gain. Our results are validated by comparison with equivalent contingent valuation values for marginal willingness to pay and marginal willingness to accept compensation for changes in job characteristics. Physicians' income preferences determine the effectiveness of 'pay for performance' and other incentive schemes. Our results may explain the relative ineffectiveness of financial incentive schemes that rely on increasing physicians' incomes.

Keywords: General Practitioners, Income, Reference-Dependence, Discrete Choice Experiment

JEL codes: I11, J44, J31

*Corresponding Author: jon.h.holte@uit.no

Acknowledgements: This project is funded by a grant from The Norwegian Research Council. We are grateful to Ivar Aaraas (GP, Professor), Peder Halvorsen (GP, PhD) and Svein Steinert (GP) for their important experience based suggestions to the chosen DCE attributes and levels, and; the 1275 busy GPs for taking time to participate in this study. When revising the paper we have benefited from the helpful comments from Anthony Scott, Marjon van der Pol, participants at NHESG meeting in Oslo 2013, and seminar participants at Monash University.

1. Introduction

An understanding of physicians' income preferences is crucial in areas of health policy that seek to affect physician behavior through financial incentives. For example, payment mechanisms in Australia aim to affect location choice of General Practitioners (GPs) through substantial financial incentives for practice in rural areas[1]. Recent pay-for performance schemes in the UK and US (Doran et al., 2006, Rosenthal and Dudley, 2007) seek to improve quality of care. The effectiveness of such schemes is in doubt (Scott et al., 2011) possibly reflecting heretofore ignored complexities in physicians preferences over changes to their income.

In the health economics literature, discrete choice experiments (DCEs) have emerged as a popular technique for eliciting preferences, often expressed in terms of marginal willingness to pay (MWTP) or marginal willingness to accept compensation (MWTA) (de Bekker-Grob et al., 2012). In a review of the application of DCEs to elicit health workers preferences for job characteristics, we find that the inferred MWTP (how much income health workers would be willing to forego for an improvement in non-pecuniary attributes) and/or MWTA (how much income health workers would need to be compensated for a deterioration in non-pecuniary attributes), is reported from seven out of eight experiments conducted in high-income countries[2], and from four out of eleven experiments conducted in low and middle income countries[3]. In a recent study on Australian GPs, Scott et al. (2013) report that the least attractive rural job package would require incentives of at least 130% of annual earnings, i.e. around AUD 237 000.

Due to the importance of DCE studies of health workers in informing policy debates around financial incentives for doctors, it is particularly important that DCEs correctly characterize health workers' income preferences. The standard approach in analysis of DCEs assumes a

[1] In Australia, the General Practice Rural Incentives Plan pays a lump sum for initial relocation, then an annual payment dependent on the years of service in the rural area. The payments are higher in 'more rural' areas. http://www.medicareaustralia.gov.au/provider/patients/rural-programs/general-practice/index.jsp

[2] (Gunther et al., 2010, Pedersen et al., 2012, Scott, 2001, Scott et al., 2013, Sivey et al., 2012, Ubach et al., 2003, Zweifel et al., 2009)

[3] (Chomitz et al., 1998, Kolstad, 2011, Rockers et al., 2012, Vujicic et al., 2011)

linear functional form for the pecuniary attribute in the respondents` utility function, i.e. that respondents have a constant marginal utility of money (Hoyos, 2010, van der Pol et al., 2014). This approach applies for all studies reporting monetary values from job choice experiments, with exception of Kolstad (2011) who use a second degree polynomial specification for income, without discussing nonlinearities in more detail. van der Pol et al. (2014) is the only study to explore non-linear functional forms in the area of health economics, although not in the setting of health workers' preferences. In a DCE eliciting public's preferences of waiting time for hip and knee replacement, they find that the linear utility function led to much higher estimates of marginal rates of substitution (willingness to wait) than the non-linear specifications, which take into account diminishing marginal utility.

The aim of this paper is to examine nonlinearity and reference-dependence in the pecuniary attribute (income) for general practitioners. We test the standard economic assumption of diminishing marginal utility. In addition, since we use '*current income'* as a level in the DCE with competing levels on each side (*less than* and *more than* current income), we are able to test the theory of reference dependent utility, where Kahneman and Tversky propose that individuals form preferences in relation to 'reference states', from which *losses loom larger than gains* (Kahneman and Tversky, 1979, Tversky and Kahneman, 1991). Furthermore, we explore the policy implications of our results for designing incentive schemes for health workers, and for general practitioners in particular.

This is the first paper to investigate reference-dependence and loss aversion in the context of job choice experiments, even though previous experiments use *current income* as a level together with positive and negative levels for income (Scott, 2001, Scott et al., 2013). Whilst reference-dependence and loss aversion in choice experiments have received attention in some recent papers within environmental and transport economics (Hess et al., 2008, Masiero and Hensher, 2010, Lanz et al., 2010), this topic is ignored in the health economics literature. Our paper also contributes to the broader literature on financial incentives for physicians and pay-for-performance (Rosenthal and Frank, 2006, Scott et al., 2011, Van Herck et al., 2010) which has largely ignored nonlinearities and reference dependence in analysing the effects of incentive schemes.

In contrast to the DCE literature, the presence of loss aversion is thoroughly examined in the empirical literature using the contingent valuation (CV) method (Horowitz and McConnell, 2002). Most studies find that the willingness to accept compensation for a loss is higher than the willingness to pay for a gain (see e.g. (Kahneman et al., 1990)). These results lend support to the theory of loss aversion, without providing evidence on the causal mechanism. Alternative theories suggest that the discrepancy can be explained by income effects (MWTP is constrained by disposable income as opposed to MWTA) and/or substitution effects (low substitutability can give rise to extreme MWTA values, which arguably can be explained within the bounds of rationality) (Hanemann, 1991).

The increasing evidence that losses in income loom larger than gains seems to be ignored in the physician job choice literature where MWTA and MWTP obtained from symmetric models (in which the ratio MWTP/MWTA = 1) are often reported interchangeably.

To test the validity of the MWTP *and* MWTA values inferred from the DCE, using the reference dependent (asymmetric) modeling approach applied in this study, we compare the results with their equivalents obtained using the contingent valuation (CV) method. This is the first convergent validity study to compare values obtained from direct and explicit MWTP *and* MWTA questions with implicit values inferred from DCE.[4] We find that the monetary values obtained using the CV and DCE method from this study are remarkably close, when we use the appropriate functional forms for the pecuniary attribute in the analysis of the DCE which account for reference-dependence.

Beyond the methodological contribution, we provide robust evidence about GPs` preferences for some key job characteristics including type of practice (private practice or salaried), control over working hours, professional development and autonomy. The paper is structured as follows. First, we explain the DCE and CV method, and discuss possible sources of bias associated with each method. Second, we provide details about how this study was designed, in attempt to minimize all known sources of biases, both for the DCE and CV method. Third and fourth, we present the results and discuss the main findings.

2. Stated preferences methods to elicit MWTA and MWTP

2.1. The discrete choice experiment (DCE)

In a discrete choice experiment respondents are presented with a series of choices between two or more alternatives. Each alternative is described alongside some selected attributes for which the levels vary systematically within a choice set. Under the assumption that respondents behave according to the axioms of random utility theory, resulting choices reveal an underlying (latent) utility function, which enables the researcher to estimate

[4] It would be preferable to use monetary values obtained from revealed preferences for comparison, but such data is not available - which usually is why stated preference methods are being applied in the first place. A close agreement between two stated choice methods (DCE and CV in this study) strengthens the validity of the methods, while a discrepancy weakens the validity, i.e. show that *at least* one method is producing erroneous estimates.

implicitly the marginal rate of substitution (MRS) between any two attributes included in the experiment (de Bekker-Grob et al., 2012). Given that a pecuniary attribute (income in our study) is included in the experiment, an implicit willingness to pay value (or willingness to accept) can be estimated for changes in all the other attributes.

The extent to which monetary values inferred from DCEs correspond to “real” MWTP or MWTA values has been challenged, as the monetary values obtained from a DCE can be sensitive to the range of the pecuniary attribute. For instance, Skjoldborg and Gyrd-Hansen (2003) find that a wider cost range including higher payments is associated with lower parameter weights associated with the payment variable, and thus lower MWTP. A crucial psychometric issue is the cognitive capacity of respondents, in terms of how many attributes with different values they can handle. Inconsistent choices may result when decision makers are faced with overly complex problems. In particular attribute non-attendance (in the form of a simplifying heuristics) is a concern. Non-attendance to an attribute results in downward bias for its coefficient, so the bias in the monetary values depends on the relative degree of bias in the pecuniary and non-pecuniary attribute (Collins, 2012).

The monetary values implied by DCEs have also been found to be sensitive to issues in the data analysis. Hole and Kolstad (2012) show that the mean MWTP differ considerably, depending on whether the monetary attribute is treated as fixed or random in the mixed logit model, which currently is the most commonly used model to analyze data from DCEs in health economics. The specification of the utility function, however, has received limited attention within the DCE literature. As stated by van der Pol et al. (2014) this lack of investigation is surprising given evidence from the contingent valuation literature suggests that welfare estimates are sensitive to different specifications of the utility function.

2.2. The contingent valuation (CV) method

In the CV method, respondents’ MWTP for a good or service (or MWTA to forego a good or service) is elicited by asking *direct* questions, in terms of an open ended format, a payment scale format, or using a dichotomous choice format. Monetary values obtained using contingent valuation questions have been questioned for a number of reasons. There is increasing evidence of MWTP values being *in*sensitive to theoretically relevant factors (e.g. size of the good), and sensitive to theoretically *ir*relevant factors (e.g. the opening bid) (Beattie et al., 1998, Bobinac et al., 2012, Herriges and Shogren, 1996, Olsen et al., 2012). Such inconsistent survey results can be explained by the hypothetical nature of contingent valuations where respondents relate to a ‘constructed market’ with hypothetical goods (Smith, 2003). However, in the current study of *experienced* GPs valuing attributes of their *own* job, the 'task familiarity' (Schläpfer and Fischhoff, 2012) of the stated preference exercise is strong.

3. Methods

3.1. Study design and sampling procedures

The study was cross-sectional, and used a structured questionnaire to collect data. All GPs in Norway (N=4305) were invited to participate through a letter sent by post May 2012. Contact information was obtained from the HELFO database, which is the same source patients use to select GPs. Respondents were informed about the purpose of the survey and asked to access a webpage to complete an online questionnaire.

The attributes chosen to be included in the DCE were*: type of practice*, *opportunity to control working hours*, *opportunity for professional development*, *degree of professional autonomy* and *income* (see Table 1). Only five attributes were selected to minimize the cognitive burden for the respondents, and thereby reducing the risk of attribute non-attendance bias. The attributes and levels were carefully chosen based on earlier studies (Scott, 2001), the ongoing MABEL study (Sivey et al., 2012, Scott et al., 2013) as well as constructive involvement of two GPs. The selected attributes and attribute levels were pilot tested in a group of GPs. Finally, the questionnaire was presented to the leaders of the Norwegian GP association who made no objections regarding any potential bias or irrelevance.

The choice sets are created in the experimental design phase. An efficient design minimizing the D-error for the multinomial logit model was created using the software Ngene provided by ChoiceMetrics (ChoiceMetrics, 2012). The design was blocked into four by minimising the average correlation between the blocking column and the attribute columns. The respondents were randomly allocated to the four blocks, and provided with five choice sets each. Figure 1 shows an example of a choice set as presented to the respondents.

Subsequent to the DCE, respondents were asked explicit open-ended questions to make partial trade-offs between income and each of the non-pecuniary job attributes included in the experiment. Through random allocation, half of the respondents were asked to state how much they would need to be compensated in terms of additional annual income (denoted $MWTA_{CV}$) for a deterioration in one attribute (e.g. for having limited rather than good opportunity to control work hours), while respondents in the other half were asked to state how much they would be willing to sacrifice in terms of annual income ($MWTP_{CV}$) for an improvement in one attribute (see Appendix 1).

3.2. Data analyses

Data from the DCE is analysed on the basis of a random utility model, assuming that respondents consider all the relevant information and always choose the alternative that yields the highest utility (McFadden, 1974). The utility doctor *i* derives from choosing alternative *j* in choice situation *t* is specified as:

$$U_{ijt} = X_{ijt}\,\beta_i + \varepsilon_{ijt},$$

$$i = 1,..,n;\ \ j = A, B, own;\ \ t = 1, \ldots, 5$$

where X_{ijt} is a vector containing the job attributes including income, β_i is a coefficient for individual GP *i* and ε is an idiosyncratic error term assumed to be extreme-value distributed.

We estimated a mixed logit model, where the coefficients for the non-pecuniary attributes were allowed to vary over individuals ($\beta_i \neq \beta$) assuming they are drawn from a normal distribution, while the parameter for income is treated as fixed ($\beta_i = \beta$). By holding the parameter for income fixed we avoid the problem of calculating MWTP and MWTA from the ratio of two normal distributions (Train, 2001, Hole and Kolstad, 2012). An alternative approach would be to specify the distribution for MWTP directly at the estimation stage, an approach known as estimation in WTP space (Hole and Kolstad, 2012). However, since the WTP space model uses the pecuniary attribute as a normalising constant, having a non-linear income attribute and/or different income parameters representing gains and losses is not desirable (Rose and Masiero, 2010). Thus, as we aim to experiment with alternative nonlinear functional forms for the monetary attribute (e.g. estimate separate coefficients for gains and losses in income) it is not appropriate to use the WTP space version of the mixed logit model. Although it may be questionable to assume that all individuals have the same income coefficients, Hole and Kolstad (2012) find that this specification (i.e. a mixed logit model with all the non-pecuniary attribute coefficients as random and the pecuniary attribute coefficient as fixed) produce means of WTP that are similar to those from the WTP space model in which the marginal utility of income is allowed to vary over individuals[5]. Finally, some literature has suggested using models which allow for scale heterogeneity in DCE analysis (Fiebig et al., 2010, Flynn et al., 2010). We have run the GMNL model for the

[5] In the same study they find that the mixed logit model with all coefficients (pecuniary and non-pecuniary) as random produce much higher (and seemingly) unrealistic means of WTP, which is in line with findings in other studies (Scarpa et al., 2008, Sonnier et al., 2007, Train and Weeks, 2005).

simpler of our specifications and find very similar results to those we present using the mixed logit model.

We use effects coding rather than dummy coding for the non-pecuniary attributes to clearly examine the effect of status quo bias (Bech and Gyrd-Hansen, 2005). For the income attribute we observe four points on the utility function: 1) NOK 100 000 *less* than current income; 2) current income; 3) NOK 150 000 *more* than current income, and; 4) NOK 300 000 *more* than current income (€1 = NOK 8). Whereas in the standard linear model, the marginal utility of income is constant, for some of the nonlinear models the effect of a change in income will depend on the absolute level of income. Where $\overline{income_i}$ is GP i's actual income[6] and $incomechange_{ijt}$ is the change in income, we create the $income_{it}$ for GP *i,* alternative *j* at choice occasion *t* as $\overline{income_i} + incomechange_{ijt}$. Using this data we run models and estimate $MWTP_{DCE}$ and $MWTA_{DCE}$ for the four non-pecuniary attributes. Seven models with alternative specifications of the income attribute are estimated. Table 3 provides an overview of the models with their respective mathematical formulae.

1) Linear income (*the standard approach*):

Our first specification is a benchmark for all the others and represents the standard approach in the literature. Income enters linearly into the utility function and implies a constant MWTP.

2) Log income:

The second specification adopts the simplest form of nonlinearity in preferences for income using the log function. This model has the advantage of parsimony, with only one coefficient to be estimated for the income variable. However, the log model imposes the specific assumption that the marginal utility of income is falling as income rises, $dU/d(income) = \frac{\beta_1}{income}$, and therefore the MWTP for the attribute Z increases as income rises.

3) Polynomial income (second degree):

The polynomial functional form, unlike the log model, allows utility to be non-monotonic in income. With the second-degree polynomial the function can have only one 'turning point' where the sign of the effect of income on utility changes. It is possible to extend the polynomial model to include higher-order terms, allowing for more than one turning point. However, we do not find support for this in our data.

4) Piecewise linear income

[6] Actual income is defined as the midpoint of the income range indicated by the responding GP in the survey.

The piecewise linear function is a direct extension of the linear utility function, allowing to estimate separate coefficients for gains and losses as we might expect for reference dependent preferences as in Kahneman and Tversky (1979). Testing for the difference between the estimate of the coefficient for income losses, β_{11}, and the estimate of the coefficient for income gains, β_{12}, will allow us to establish if income preferences are reference-dependent, conditional on the linear functional form assumption.

5) Piecewise log income

The piecewise linear model described above could estimate different coefficients for losses and gains simply as a result of the true utility function being nonlinear, but not reference-dependent. Therefore we can also estimate a nonlinear functional form, the log function, *and* allow for reference-dependent preferences by creating a spline for log income in the same way as for linear income. This model will help us distinguish between nonlinearities in the functional form and reference dependence itself.

Testing for the difference between the estimate of the coefficient for income losses, β_{11}, and the estimate of the coefficient for income gains, β_{12}, will allow us to establish if income preferences are reference-dependent, conditional on the log functional form assumption for income.

6) Piecewise polynomial income

In a similar way, we can also estimate the alternative nonlinear functional form, the polynomial income model, whilst allowing for reference-dependence. This model will help establish if there is reference-dependence in addition to nonlinearities that can be captured by a squared income term.

Again, a test of the difference between the estimates of β_{11} and β_{12} will test if income preferences are reference-dependent, conditional on the second degree polynomial specification. Note that we only include reference-dependence in the first (linear) term of the polynomial. With only four levels of income (and hence three 'changes' in income) we can identify a maximum of three income parameters, in this case β_{11}, β_{12} and β_{21}.

7) Effects coding:

This model is the opposite of the linear and log models in that it estimates the most flexible possible form for income preferences (given our DCE) and requires the estimation of the highest number of parameters. This model estimates a different utility level and monetary value for each level of income. As each effects coded income variable is scaled to represent a particular change in income, we need to multiply by that level of income to obtain the MWTA and MWTP in money terms.

The effects coded model can be used to assess the validity of the assumptions in the other models. For example from the estimates of the effects-coded model it will be possible to assess, and test formally, whether the linear model or log models are a good approximation of the preference structure. We will be able to use the estimates to test if marginal utility is constant, $MWTP^{11} = MWTA^{12} = MWTA^{13}$ as implied by the linear model, and also if marginal utility is decreasing in income $MWTP^{11} < MWTA^{12} < MWTA^{13}$, as implied by the log model.

To ensure comparable figures, we include only respondents who provide *valid* answers to all the CV and DCE questions in the analysis. Respondents with missing value or ambiguous answer for any question, zero values for all open ended CV questions, and a $MWTP_{CV}$ or $MWTA_{CV}$ higher than current income were excluded from the analysis.

4. Results

4.1 Sampling characteristics

A total of 1275 (30%) GPs answered the survey, out of which 934 were considered eligible in according to the inclusion criteria.[7] The respondents are largely representative of Norwegian GPs according to age, gender, number of listed patients and geographical distribution, while specialists in general medicine are overrepresented in our sample (see Table 2). The mean income, constructed from the mid-points of the selected income range for each GP, is approximately NOK 1 050 000 (SD≈ NOK 298 000).

4.2 Comparing models

Table 4 presents results from the mixed logit model with different specifications of the income variable in the utility function. Reviewing the summary statistics, we find that the model with *effects coding* for income (without parametric assumptions) fits the data best according to the Bayesian Information Criterion (BIC[8]). The model with linear income (which is standard in the

[7] 10 respondents were excluded since they no longer work as a GP, 156 respondents due to missing values on CV questions (the DCE questions were obligatory), 99 respondents due ambiguous answers and refusals such as \`find it difficult to answer\` and \`not willing to answer\`, 42 respondents due to zero values on all CV questions, 8 respondents due to missing information about current income and 26 respondents since they provide MWTP or MWTA above current income in the CV questions - a few of those report very high MWTA (up to 12 digits) which substantially inflate the mean estimates.
[8] The BIC is calculated as follows: $BIC = -2lnL + kln(n)$, where lnL is the log-liklihood, k is the number of parameters, and n is the number of observations of the estimated model.

prior literature) is clearly outperformed by the model with effects coding in according to BIC, suggesting that we should look for alternative utility specifications. The model with polynomial income fits the data slightly better compared to the model with linear income, while the model with log income fits the data worse. All the models with piecewise (asymmetric) specifications of income fit the data substantially better than the models with symmetric specifications, of which the piecewise linear income specification performs best.

4.3 Nonlinearity and reference dependence

Turning to the coefficients for income we find that losses loom larger than gains. All three models with piecewise income specifications as well as the effects coded specification predict much larger changes in utility associated with reductions in income than the equivalent increases in income.[9] This asymmetry is illustrated in Figure 2, where the plot for effects coded income show that the reduction in utility of moving from current income to *NOK 100 000 less* is approximately of equal size to the increase in utility generated by a shift from current income to *NOK 300 000 more*, i.e. a three-fold difference. For the piecewise specifications of income, the asymmetry is illustrated clearly by the kink at the point of current income. It can be seen that the symmetric linear specification of income overestimates the positive utility effects of increases in income and underestimates the negative utility effects of decreases in income.

The coefficients for *effects coded income* indicate that the marginal utility of gains in income decreases substantially for higher levels, i.e. a change from current income to NOK 300 000 *more* is valued only 1.3 times higher than a change from current income to NOK 150 000 more, despite the two-fold difference in money value (see Figure 2). This is captured (to some extent) in the polynomial utility specifications, by the negative and statistically significant coefficient for income squared.

4.4. Comparing MWTP and MWTA

Table 5 presents mean monetary values obtained using different utility specifications of income, in addition to the $MWTP_{CV}$ and $MWTA_{CV}$ obtained using the CV method. We find that loss aversion in the income attribute give rise to a substantial degree of $MWTP_{DCE}$ – $MWTA_{DCE}$ asymmetry, evident for all utility specifications of income, i.e. $MWTP_{DCE}$ <

[9] The difference between the estimate of the coefficient for income losses, β_{11}, and the estimate of the coefficient for income gains, β_{12} is statistically significant in all the models with piecewise specifications.
Piecewise linear: $\beta_{11} = 23.3$; 95%CI=19.7 - 26.8, $\beta_{12} = 5.9$; 95%CI=5.0 - 6.8
Piecewise log: $\beta_{11} = 18.1$; 95%CI=15.1 - 21.1 $\beta_{12} = 6.5$; 95%CI=5.5 - 7.4
Piecewise polynomial: $\beta_{11} = 24.9$; 95%CI=21.0 - 28.8 $\beta_{12} = 8.2$; 95%CI=5.9 - 10.5

$MWTA_{DCE}$. The $MWTA_{DCE}/MWTP_{DCE}$ ratio equals 3.9, estimated on the basis of the model with piecewise linear income (which provide the best fit for our data).[10]

The results on $MWTP_{DCE}$ – $MWTA_{DCE}$ asymmetry are echoed by results from the CV method. On the basis of the model with linear piecewise specification for income, we find that the $MWTP_{DCE}/MWTP_{CV}$ ratio is 0.9, 1.1 and 0.8 for ´hours`, ´prof development` and ´autonomy` respectively, while the corresponding $MWTA_{DCE}/MWTA_{CV}$ ratio is 1.4, 1.3 and 1.1.

5. Discussion

This paper has examined nonlinearity and reference-dependence in GPs income preferences in the context of a DCE. We find that the symmetric linear specification of income, standard in previous applied studies, overestimates the positive utility effects of increases in income and underestimates the negative utility effects of decreases in income. This finding of reference-dependence is not because of, but is in addition to, nonlinearities in the income effect (i.e. a diminishing marginal utility of income). The presence of loss aversion in the income attribute gives rise to a substantial degree of MWTP-MWTA asymmetry. These results are echoed by results from some recent studies in the area of environment and transport (Hess et al., 2008, Masiero and Hensher, 2010, Lanz et al., 2010). Furthermore, our findings conform with results from previous (non-experimental) studies, which suggest that the reference- or target-income hypothesis is applicable to physicians, without quantifying the extent to which loss aversion in income is a source to observed MWTA-MWTP asymmetry (Rizzo and Zeckhauser, 2003, Rizzo and Blumenthal, 1996).

This study has significant policy implications as doctors' income preferences determine the effects of financial incentives present in many areas of health policy. In particular, the study suggests that it may be more effective to design financial incentives as penalties, so that GPs need to respond to avoid *losses* in income, rather than using bonuses implemented on top of their current income. Using penalties would probably be more controversial – and certainly less acceptable among GPs than using rewards, which potentially may have negative effects, e.g. cause opportunistic behavior (Eijkenaar, 2013). This is claimed to be the reason why pay for performance programs typically only provide positive incentives (Eijkenaar, 2013). Nevertheless, the findings from this study could explain the relatively small impacts of large financial incentive schemes such as the Quality and Outcomes Framework in the UK (Campbell et al., 2009).

[10] Since $MWTA_{DCE}$ – $MWTP_{DCE}$ asymmetry in our estimates only hinges upon loss aversion in the income attribute, the degree of asymmetry is equal for all the attributes.

In terms of the DCE literature, we encourage choice modellers to investigate nonlinearity in preferences over DCE attributes, as failure to do so potentially can bias estimates substantially and mask policy relevant information. This is particularly evident for all experiments that include attribute levels which can be perceived as ´references states`, together with positive and negative levels. Note that certain levels potentially can be perceived as ´references states` even though they are not explicitly anchored to respondents` current level (as in the present study), e.g. in experiments using absolute levels of income around the average starting salary, which is common in job choice experiments including students.

The MWTP and MWTA inferred from the DCE seem plausible in absolute and relative values, and are further supported by the close agreement with monetary values obtained from the CV questions. This is the first convergent validity study to compare implicit values (for each attribute) inferred from DCE with values obtained from direct and explicit MWTP *and* MWTA questions. We find that the MWTA and MWTP obtained using the CV and DCE method from this study correspond closely, although $MWTA_{DCE}$ is consistently higher than $MWTA_{CV}$, with a ratio in the range from 1.1 – 1.4. The same pattern (i.e. that monetary values derived from the DCE tend to be higher than those from the contingent valuation method) is found in two other studies that compare MWTP derived from the open-ended contingent valuation method and the DCE method (Bijlenga et al., 2011, van der Pol et al., 2008). In a study of lay-persons' MWTP in obstetrics, Bijlenga et al. (2011) find that the implied MWTP derived with DCE was between 2.3 and 10.2 times higher than with CV. In a study of pregnant women`s MWTP for prenatal care, van der Pol et al. (2008) find that MWTP values produced by the DCE were about twice as high as the mean estimates produced by the open ended question. More surprisingly, we find that $MWTP_{DCE}$ is lower than $MWTP_{CV}$ for `hours` and `autonomy`. A possible explanation is that we ask about MWTP and MWTA for each attribute, as opposed to previous convergent validity studies, all of which ask about MWTP for a *total* package. Nevertheless, the magnitude of the discrepancy is substantially smaller in our study as compared to previous convergent validity studies.

A possible explanation for the relatively close agreement between the CV and DCE method is that the respondents (doctors) in this study have more well informed preferences for the attributes in question (i.e. job attributes which they know through years of experience), as compared to lay-persons and patients in previous studies. Evidence from a recent DCE of Norwegian doctors (GPs and hospital consultants) that examines prescribing behaviour, suggest that the attribute *non-attendance prevalence* is low (Hole et al., 2013). Furthermore, the doctors in our study may well find it easier to answer the direct CV questions than the lay persons and patients in previous convergent validity studies, since trade-offs between work and income essentially is the basis of the labour market. For lay persons and patients it may be more difficult to assign monetary values for non-marketed goods such as health services.

The policy context of the current study is one where the mean income of GPs appears to be at least as high as the mean income of hospital doctors. In many other countries GPs earn substantially less than specialist hospital doctors (Bodenheimer et al., 2007, Cheng et al., 2012). GPs who earn less than specialists may find the difference in wages as unfair and therefore consider the wage level of specialists as their reference point instead of their *current income*, so that gains in income up to the wage level for specialists is perceived as *reductions in losses* rather than gain (Bateman et al., 2005). It remains to test if increases in income have a larger impact on utility and job choices in countries where GPs earn substantially less than specialists. The theory of reference dependent preferences is silent about how reference states are being determined, although a natural interpretation is that an individual's reference state is the bundle of goods she currently owns (Bateman et al., 2005).

Aiming at informing policy, this study was designed in line with previous job choice experiments, using similar attributes and levels. The advantage is that it allows us to examine the effect of not taking nonlinearity into account in standard applied studies, i.e. in studies using an easily manageable number of levels that are assumed to be realistic for the respondents and relevant for policy purposes. For the purpose of investigating loss aversion as social phenomenon it would be an advantage to include more levels for the income attribute, in particular use more than one level in the loss domain to examine whether marginal utility decreases at a different rate for losses and gains. For the model with a polynomial piecewise specification of income we simply assume that marginal utility decreases at a similar rate. This should be examined in future studies.

An additional argument for using reference pivoted designs is that it may ease the cognitive burden for respondents, since individuals typically evaluate possible alternatives (e.g. jobs) on the basis of their *current* situation. Thus, although reference dependence and loss aversion to a large extent have been ignored in the area of health economics, as opposed to other areas of research, there may be good reasons for using reference pivoted designs in future studies, in various study contexts.

Acknowledgements

This project is funded by a grant from The Norwegian Research Council. We are grateful to Ivar Aaraas (GP, Professor), Peder Halvorsen (GP, PhD) and Svein Steinert (GP) for their important experience based suggestions to the chosen DCE attributes and levels, and; the 1275 busy GPs for taking time to participate in this study. When revising the paper we have benefited from the helpful comments from Anthony Scott, Marjon van der Pol, participants at NHESG meeting in Oslo 2013, and seminar participants at Monash University.

BATEMAN, I., KAHNEMAN, D., MUNRO, A., STARMER, C. & SUGDEN, R. 2005. Testing competing models of loss aversion: an adversarial collaboration. *Journal of Public Economics* 1561-1580.

BEATTIE, J., COVEY, J., DOLAN, P., HOPKINS, L., JONES-LEE, M., LOOMES, G., PIDGEON, N., ROBINSON, A. & SPENCER, A. 1998. On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 1-Caveat Investigator. *Journal of Risk and Uncertainty,* 17**,** 5-26.

BECH, M. & GYRD-HANSEN, D. 2005. Effects coding in discrete choice experiments. *Health Economics,* 14**,** 1079-1083.

BIJLENGA, D., BONSEL, G. J. & BIRNIE, E. 2011. Eliciting willingness to pay in obstetrics: comparing a direct and an indirect valuation method for complex health outcomes. *Health Economics,* 20**,** 1392-1406.

BOBINAC, A., VAN EXEL, N. J. A., RUTTEN, F. F. H. & BROUWER, W. B. F. 2012. GET MORE, PAY MORE? An elaborate test of construct validity of willingness to pay per QALY estimates obtained through contingent valuation. *J Health Econ,* 31**,** 158-168.

BODENHEIMER, T., BERENSON, R. A. & RUDOLF, P. 2007. The Primary Care–Specialty Income Gap: Why It Matters. *Annals of Internal Medicine,* 146**,** 301-306.

CAMPBELL, S. M., REEVES, D., KONTOPANTELIS, E., SIBBALD, B. & ROLAND, M. 2009. Effects of Pay for Performance on the Quality of Primary Care in England. *New England Journal of Medicine,* 361**,** 368-378.

CHENG, T. C., SCOTT, A., JEON, S.-H., KALB, G., HUMPHREYS, J. & JOYCE, C. 2012. WHAT FACTORS INFLUENCE THE EARNINGS OF GENERAL PRACTITIONERS AND MEDICAL SPECIALISTS? EVIDENCE FROM THE MEDICINE IN AUSTRALIA: BALANCING EMPLOYMENT AND LIFE SURVEY. *Health Economics,* 21**,** 1300-1317.

CHOICEMETRICS 2012. Ngene 1.1.1 User Manual & Reference Guide, Australia.

CHOMITZ, K., SETIADI, G., AZWAR, A. & WIDIYARTI, I. N. 1998. What do doctors want? Developing incentives for doctors to serve in Indonesia's rural and remote areas. *World Bank Policy Research Working Paper, Washington, DC,* #1888.

COLLINS, A. T. 2012. Attribute nonattendance in discrete choice models: measurement of bias, and a model for the inference of both nonattendance and taste heterogeneity', PhD Thesis, The University of Sydney, Sydney, NSW, Australia.

DE BEKKER-GROB, E. W., RYAN, M. & GERARD, K. 2012. Discrete choice experiments in health economics: a review of the literature. *Health Economics,* 21**,** 145-172.

DORAN, T., FULLWOOD, C., GRAVELLE, H., REEVES, D., KONTOPANTELIS, E., HIROEH, U. & ROLAND, M. 2006. Pay-for-Performance Programs in Family Practices in the United Kingdom. *New England Journal of Medicine,* 355**,** 375-384.

EIJKENAAR, F. 2013. Key issues in the design of pay for performance programs. *The European Journal of Health Economics,* 14**,** 117-131.

FIEBIG, D. G., KEANE, M. P., LOUVIERE, J. & WASI, N. 2010. The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity. *Marketing Science,* 29**,** 393-421.

FLYNN, T. N., LOUVIERE, J. J., PETERS, T. J. & COAST, J. 2010. Using discrete choice experiments to understand preferences for quality of life. Variance-scale heterogeneity matters. *Social Science & Medicine,* 70**,** 1957-1965.

GUNTHER, O. H., KURSTEIN, B., RIEDEL-HELLER, S. G. & KONIG, H. H. 2010. The Role of Monetary and Nonmonetary Incentives on the Choice of Practice Establishment: A Stated Preference Study of Young Physicians in Germany. *Health Services Research,* 45**,** 212-229.

HANEMANN, W. M. 1991. Willingness to Pay and Willingness to Accept: How Much Can They Differ? *The American Economic Review,* 81**,** 635-647.

HERRIGES, J. A. & SHOGREN, J. F. 1996. Starting Point Bias in Dichotomous Choice Valuation with Follow-Up Questioning. *Journal of Environmental Economics and Management,* 30**,** 112-131.

HESS, S., ROSE, J. M. & HENSHER, D. A. 2008. Asymmetric preference formation in willingness to pay estimates in discrete choice models. *Transportation Research Part E: Logistics and Transportation Review,* 44**,** 847-863.

HOLE, A. R. & KOLSTAD, J. R. 2012. Mixed logit estimation of willingness to pay distributions: a comparison of models in preference and WTP space using data from a health-related choice experiment. *Empirical Economics,* 42**,** 445-469.

HOLE, A. R., KOLSTAD, J. R. & GYRD-HANSEN, D. 2013. Inferred vs. stated attribute non-attendance in choice experiments: A study of doctors' prescription behaviour. *Journal of Economic Behavior & Organization,* 96**,** 21-31.

HOROWITZ, J. K. & MCCONNELL, K. E. 2002. A review of WTA/WTP studies. *Journal of Environmental Economics and Management,* 44**,** 426-447.

HOYOS, D. 2010. The state of the art of environmental valuation with discrete choice experiments. *Ecological Economics,* 69**,** 1595-1603.

KAHNEMAN, D., KNETSCH, J. L. & THALER, R. H. 1990. Experimental Tests of the Endowment Effect and the Coase Theorem. *Journal of Political Economy,* 98**,** 1325-1348.

KAHNEMAN, D. & TVERSKY, A. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica,* 47**,** 263-291.

KOLSTAD, J. R. 2011. HOW TO MAKE RURAL JOBS MORE ATTRACTIVE TO HEALTH WORKERS. FINDINGS FROM A DISCRETE CHOICE EXPERIMENT IN TANZANIA. *Health Economics,* 20**,** 196-211.

LANZ, B., PROVINS, B., BATEMAN, I. J., SCARPA, R., WILLIS, K. & OZDERMIROGLU, E. 2010. Investigating Willingness to Pay – Willingness to Accept Asymmetry in Choice Experiments. *Choice Modelling The State-of-the-art and the State-of-practice - Proceedings from the Inaugural International Choice Modelling Conference,* Chapter 24.

MASIERO, L. & HENSHER, D. A. 2010. Analyzing loss aversion and diminishing sensitivity in a freight transport stated choice experiment. *Transportation Research Part A: Policy and Practice,* 44**,** 349-358.

MCFADDEN, D. 1974. *Conditional Logit Analysis of Qualitative Choice Behaviour, in Frontiers in Econometrics***,** 105-142.

OLSEN, J. A., RØGEBERG, O. J. & STAVEM, K. 2012. What explains willingness-to-pay for smoking-cessation treatments: Addiction level, quit rate effectiveness or the opening bid? *Appl Health Econ Health Policy,* 10**,** 407–15.

PEDERSEN, L. B., KJAER, T., KRAGSTRUP, J. & GYRD-HANSEN, D. 2012. General practitioners' preferences for the organisation of primary care: A discrete choice experiment. *Health Policy,* 106**,** 246-256.

RIZZO, J. A. & BLUMENTHAL, D. 1996. Is the Target Income Hypothesis an Economic Heresy? *Medical Care Research and Review,* 53**,** 243-266.

RIZZO, J. A. & ZECKHAUSER, R. J. 2003. Reference Incomes, Loss Aversion, and Physician Behavior. *Review of Economics and Statistics,* 85**,** 909-922.

ROCKERS, P. C., JASKIEWICZ, W., WURTS, L., KRUK, M. E., MGOMELLA, G. S., NTALAZI, F. & TULENKO, K. 2012. Preferences for working in rural clinics among trainee health professionals in Uganda: a discrete choice experiment. *Bmc Health Services Research,* 12.

ROSE, J. M. & MASIERO, L. 2010. A comparison of the impacts of aspects of prospect theory on WTP/WTA estimated in preference and WTP/WTA space. *European Journal of Transport and Infrastructure Research,* 10**,** 330-346.

ROSENTHAL, M. B. & DUDLEY, R. 2007. Pay-for-performance: Will the latest payment trend improve care? *JAMA,* 297**,** 740-744.
ROSENTHAL, M. B. & FRANK, R. G. 2006. What Is the Empirical Basis for Paying for Quality in Health Care? *Medical Care Research and Review,* 63**,** 135-157.
SCARPA, R., THIENE, M. & TRAIN, K. 2008. Utility in Willingness to Pay Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps. *American Journal of Agricultural Economics,* 90**,** 994-1010.
SCHLÄPFER, F. & FISCHHOFF, B. 2012. Task familiarity and contextual cues predict hypothetical bias in a meta-analysis of stated preference studies. *Ecological Economics,* 81**,** 44-47.
SCOTT, A. 2001. Eliciting GPs' preferences for pecuniary and non-pecuniary job characteristics. *Journal of Health Economics,* 20**,** 329-347.
SCOTT, A., SIVEY, P., AIT, O. D., WILLENBERG, L., NACCARELLA, L., FURLER, J. & YOUNG, D. 2011. The effect of financial incentives on the quality of health care provided by primary care physicians. *Cochrane Database of Systematic Reviews*.
SCOTT, A., WITT, J., HUMPHREYS, J., CATHERINE, J., GUYONNE, K., JEON, H. & MCGRAIL, M. 2013. Getting Doctors into the Bush: General Practitioners' Preferences for Rural Location. *Social Science & Medicine,* Forthcoming.
SIVEY, P., SCOTT, A., WITT, J., JOYCE, C. & HUMPHREYS, J. 2012. Junior doctors' preferences for specialty choice. *Journal of Health Economics,* 31**,** 813-823.
SKJOLDBORG, U. & GYRD-HANSEN, D. 2003. Conjoint analysis. The cost variable: an Achilles' heel? *Health Economics***,** 479 - 492.
SMITH, R. D. 2003. Construction of the contingent valuation market in health care:a critical assessment. *Health Economics,* 12**,** 609-628.
SONNIER, G., AINSLIE, A. & OTTER, T. 2007. Heterogeneity distributions of willingness-to-pay in choice models. *Quantitative Marketing and Economics,* 5**,** 313-331.
TRAIN, K. 2001. A comparison of hierarchical Bayes and maximum simulated likelihood for mixed logit. University of California, Berkeley 1-13.
TRAIN, K. & WEEKS, M. 2005. Discrete Choice Models in Preference Space and Willingness-to-Pay Space. *In:* SCARPA, R. & ALBERINI, A. (eds.) *Applications of Simulation Methods in Environmental and Resource Economics.* Springer Netherlands.
TVERSKY, A. & KAHNEMAN, D. 1991. Loss Aversion in Riskless Choice: A Reference-Dependent Model. *The Quarterly Journal of Economics,* 106**,** 1039-1061.
UBACH, C., SCOTT, A., FRENCH, F., AWRAMENKO, M. & NEEDHAM, G. 2003. What do hospital consultants value about their jobs? A discrete choice experiment. *British Medical Journal,* 326**,** 1432-1435.
VAN DER POL, M., CURRIE, G., KROMM, S. & RYAN, M. 2014. Specification of the utility function in Discrete Choice Experiments. *Value in Health,* In press.
VAN DER POL, M., SHIELL, A., AU, F., JOHNSTON, D. & TOUGH, S. 2008. Convergent validity between a discrete choice experiment and a direct, open-ended method: Comparison of preferred attribute levels and willingness to pay estimates. *Social Science & Medicine***,** 2043-2050.
VAN HERCK, P., DE SMEDT, D., ANNEMANS, L., REMMEN, R., ROSENTHAL, M. & SERMEUS, W. 2010. Systematic review: Effects, design choices, and context of pay-for-performance in health care. *Bmc Health Services Research,* 10**,** 247.
VUJICIC, M., SHENGELIA, B., ALFANO, M. & THU, H. B. 2011. Physician shortages in rural Vietnam: Using a labor market approach to inform policy. *Social Science & Medicine,* 73**,** 970-977.
ZWEIFEL, P., RISCHATSCH, M. & BRÄNDLE, A. 2009. GPs' preferences: What price fee-for-service. *Working paper series, Socioeconomic Institute, University of Zurich*.

Table 1 Attributes and levels

Attributes	**Coding**	**Levels**	**Expected sign**
Type of practice	Effects	• Fixed salary (reference level) • Private practice	?
Opportunity to control working hours	Effects	• Limited (reference level) • Very good	+
Opportunity for professional development	Effects	• Limited (reference level) • Very good	+
Degree of professional autonomy	Effects	• Limited (reference level) • High degree	+
Income	Various	• NOK 100 000 *less* than current income • Current income • NOK 150 000 *more* than current income • NOK 300 000 *more* than current income	- + +

Table 2 Characteristics

Variable	Respondents	All Norwegian GPs
Age (N=896)	47	49 [2]
Females (N=934)	37 %	35 %[2]
Number of patients listed (N=918)	1139	1182[2]
Municipality (N=934)		
<5000	16%	14 %[1]
5000-49999	53%	52 %[1]
50000 +	31%	34 %[1]
Specialty attainment (N=934)	70 %	55 %[3]
Categorical income (N=934)		
< NOK 700 000	10%	-
NOK 700 000 – 849 000	17%	-
NOK 850 000 – 999 000	20%	-
NOK 1 000 000 – 1 149 000	17%	-
NOK 1 150 000 – 1 299 000	18%	-
NOK 1 300 000 – 1 500 000	10%	-
> NOK 1 500 000	8%	-
Mean income (SD)[a]	1 051 000 (298 000)	-

1) Statistics Norway (www.ssb.no accessed 24th of March 2011)
2) http://www.helsedirektoratet.no/refusjonsordninger/tall_analyse/fastlege/flere_fastleger_i_2009__781144
3) http://www.legeforeningen.no/id/18 14.04.2011

Note: Right column is copied from (Halvorsen et al., 2012)

- Officially data on GPs income is not available [a] Mean income is constructed from the mid-points of the selected income range for each GP

Table 3 The specifications of the seven models

Specification	Model	Estimation of MWTP and MWTA
Linear income	$U = \beta_1 income + \beta_2 Z + \varepsilon$	$MWTP_Z = MWTA_Z = \frac{dU/dZ}{dU/d(income)} = \frac{\beta_2}{\beta_1}$
Log income	$U = \beta_1 \ln(income) + \beta_2 Z + \varepsilon$	$MWTP_Z = MWTA_Z = \frac{dU/dZ}{dU/d(income)} = \left(\frac{\beta_2}{\beta_1}\right) * income$
Polynomial income (second degree)	$U = \beta_1 income + \beta_{12} income^2 + \beta_2 Z + \varepsilon$	$MWTP_Z = MWTA_Z = \frac{dU/dZ}{\frac{dU}{d(income)}} = \frac{\beta_2}{(\beta_1 + 2\beta_{12} income)}$
Piecewise linear[a]	$U = [income < \overline{income}]\beta_{11} income + [income \geq \overline{income}](\beta_{11}\overline{income} + \beta_{12}(income - \overline{income}) + \beta 2 * Z + \varepsilon$	$MWTP_Z = \frac{dU/dZ}{\frac{dU}{d(income)}} = \frac{\beta_2}{\beta_{11}}, \quad income < \overline{income}$ $MWTA_Z = \frac{dU/dZ}{\frac{dU}{d(income)}} = \frac{\beta_2}{\beta_{12}}, \quad income \geq \overline{income}$
Piecewise log[a]	$U = [income < \overline{income}]\beta_{11}\ln(income) + [income \geq \overline{income}]\left(\beta_{11}\ln(\overline{income}) + \beta_{12}(ln(income) - \ln(\overline{income}))\right) + \beta_2 Z + \varepsilon$	$MWTP_Z = \frac{dU/dZ}{\frac{dU}{d(income)}} = \left(\frac{\beta_2}{\beta_{11}}\right) income, \quad income < \overline{income}$ $MWTA_Z = \frac{dU/dZ}{\frac{dU}{d(income)}} = \left(\frac{\beta_2}{\beta_{12}}\right) income, \quad income \geq \overline{income}$
Piecewise polynomial[a]	$U = ([income < \overline{income}]\beta_{11}\text{income} + \beta_{21} income^2) + [income \geq \overline{income}](\beta_{11}\overline{income} + \beta_{12}(\text{income} - \overline{income}) + \beta_{21} income^2) + \beta_2 Z + \varepsilon]$	$MWTP_Z = \frac{dU/dZ}{\frac{dU}{d(income)}} = \frac{\beta_2}{(\beta_{11} + 2\beta_{21} income)}, \quad income < \overline{income}$ $MWTA_Z = \frac{dU/dZ}{\frac{dU}{d(income)}} = \frac{\beta_2}{(\beta_{12} + 2\beta_{21} income)}, \quad income \geq \overline{income}$
Effects coding	$U = \beta_{11} minus100 + \beta_{12} plus150 + \beta_{13} plus300 + \beta_2 Z + \varepsilon$	$MWTP_Z^{11} = \frac{dU/dZ}{dU/d(minus100)} * 100{,}000 = \left(\frac{\beta_2}{\beta_{11}}\right) * -100{,}000$ $MWTA_Z^{12} = \frac{dU/dZ}{dU/d(plus150)} * 150{,}000 = \left(\frac{\beta_2}{\beta_{12}}\right) * 150{,}000$ $MWTA_Z^{13} = \frac{dU/dZ}{dU/d(plus300)} * 300{,}000 = \left(\frac{\beta_2}{\beta_{13}}\right) * 300{,}000$

Note: In the equations in this table we simplify the exposition of the alternative models in four different ways. First, we include only one other representative attribute apart from income, denoted *Z,* where *Z* represents one of the attributes in *X*. Second, we omit subscripts for the variables $U_{ijt}, income_{ijt}, Z_{ijt}, \varepsilon_{ijt}, \overline{income_i}, minus100_{ijt}, plus150_{ijt}, plus300_{ijt}$ Third, we write the equations as if both income and *Z* are continuous variables. In reality all variables apart from income are dichotomous, which slightly complicates the calculation of marginal utilities. However, treating the variable as continuous is a close approximation of the true calculation. Fourth, we write the equations as if they are simple logit models rather than mixed logit models (we omit the 'i' subscript on the β's), and therefore write the WTP values as if they are single estimates rather than normally distributed random variables. [a] For the piecewise utility specifications we create a spline for the income attribute allowing for different coefficients either side of the respondents 'actual' income denoted by $\overline{income}$. In this specification square brackets are indicator variables for the inequality within them.

Table 4 Main model, with various specifications of income

	Linear		**Log**		**Polynomial**		**Piecewise linear**		**Piecewise log**		**PCW polynomial**		**Effects coding**	
BIC	6 704.632		6 720.485		6 696.634		6 617.138		6 670.690		6 622.113		6 601.772	
Log-Likelihood	-3 290.26		-3 298.18		-3 281.48		-3 241.74		-3 268.51		-3 239.45		-3 229.28	
Number of obs.	14 010		14 010		14 010		14 010		14 010		14 010		14 010	
Variables	Mean coef/se	SD coef/se	Mean coef/se	SD coef/se	Mean coef/se	SD coef/se	Mean coef/se	SD coef/se	Mean coef/se	SD coef/se	Mean coef/se	SD coef/se	Mean coef/se	SD coef/se
Income (symetric)	8.403*** (0.410)		8.451*** (0.417)		12.064*** (1.115)									
Income losses							23.272*** (1.808)		18.124*** (1.519)		24.914*** (1.985)		-2.400***[a] (0.139)	
Income gains							5.927*** (0.450)		6.469*** (0.472)		8.225*** (1.180)		1.150***[b] (0.088)	
Income squared					-1.58E-06*** (4.33E-07)						-9.53E-07** (4.46E-07)		1.586***[c] (0.091)	
Type of practice	-0.529*** (0.057)	0.957*** (0.068)	-0.532*** (0.056)	0.937*** (0.067)	-0.536*** (0.057)	0.939*** (0.065)	-0.518*** (0.059)	1.023*** (0.069)	-0.522*** (0.058)	0.991*** (0.067)	-0.521*** (0.059)	1.017*** (0.069)	-0.503*** (0.060)	1.041*** (0.070)
Hours	-1.080*** (0.065)	0.787*** (0.073)	-1.058*** (0.064)	0.790*** (0.072)	-1.081*** (0.065)	0.787*** (0.073)	-1.108*** (0.066)	0.784*** (0.073)	-1.086*** (0.065)	0.795*** (0.072)	-1.102*** (0.066)	0.785*** (0.073)	-1.130*** (0.068)	0.801*** (0.074)
Prof. development	-0.799*** (0.054)	0.591*** (0.071)	-0.786*** (0.053)	0.579*** (0.071)	-0.806*** (0.055)	0.599*** (0.070)	-0.890*** (0.058)	0.608*** (0.071)	-0.861*** (0.057)	0.605*** (0.069)	-0.883*** (0.058)	0.608*** (0.070)	-0.891*** (0.060)	0.637*** (0.072)
Autonomy	-0.878*** (0.068)	0.868*** (0.077)	-0.860*** (0.067)	0.869*** (0.076)	-0.881*** (0.068)	0.878*** (0.076)	-0.851*** (0.068)	0.870*** (0.078)	-0.848*** (0.067)	0.865*** (0.076)	-0.847*** (0.068)	0.869*** (0.077)	-0.892*** (0.070)	0.885*** (0.078)
ASCA	-1.645*** (0.110)	0.778*** (0.156)	-1.529*** (0.106)	0.743*** (0.158)	-1.635*** (0.111)	0.763*** (0.162)	-1.046*** (0.123)	0.822*** (0.161)	-1.092*** (0.120)	0.754*** (0.165)	-1.061*** (0.123)	0.812*** (0.162)	-1.287*** (0.137)	0.819*** (0.165)
ASCB	-1.936*** (0.124)	1.002*** (0.147)	-1.817*** (0.120)	0.964*** (0.149)	-1.918*** (0.124)	0.992*** (0.150)	-1.336*** (0.133)	1.009*** (0.154)	-1.379*** (0.130)	0.968*** (0.152)	-1.351*** (0.133)	0.999*** (0.155)	-1.575*** (0.147)	0.991*** (0.157)

*** p<0.01, ** p<0.05, * p<0.1 Note: All the coefficients for income are rescaled in the model to ease interpretation (i.e. the income variables are divided with 1000000 before estimating the model to avoid extremely small numbers) [a] Coefficient for income ` NOK 100 000 less` [b] Coefficient for income ` NOK 150 000 more` [c] Coefficient for income ` NOK 300 000 more` [a-c] In the model with effects coding for income the marginal utility is not simply given by the estimated model coefficients, as in the other models. For effect coded attributes, the marginal utility is defined as the difference between the respective attribute level and the reference level (Pedersen et al., 2012, Scott et al., 2013). The reference levels are calculated by summing the other levels and multiplying by minus one, i.e. for income the reference level is -0.336.Thus, the coefficients for marginal utility, representing the change in utility of moving from the reference level (current income) to NOK 100 000 *less*, NOK 150 000 and NOK 300 000 *more* are -2.064, 1.486 and 1.922 respectively, (these are the coefficients we use to estimate monetary values).

Table 5 The monetary valuation of each attributes in each model

Model	Linear	Log	Polynomial	Piecewise linear		Piecewise log		PCW polynomial		Effects coding			CV	
	MWTP=MWTA	MWTP=MWTA	MWTP=MWTA	MWTP	MWTA	MWTP	MWTA	MWTP	MWTA	MWTP[11]	MWTA[12]	MWTA[13]	MWTP	MWTA
Type of practice	126000	132000	103000	45000	175000	61000	170000	44000	144000	49000	102000	157000	-	-
Control working hours	257000	263000	208000	95000	374000	126000	353000	92000	306000	110000	228000	353000	101000	268000
Prof. development	190000	196000	155000	76000	300000	100000	280000	74000	245000	86000	180000	278000	71000	224000
Degree of autonomy	209000	214000	170000	73000	287000	98000	276000	71000	235000	86000	180000	278000	92000	254000

- CV questions not applicable for `type of practice`

HALVORSEN, P. A., STEINERT, S. & AARAAS, I. J. 2012. Remuneration and organization in general practice: Do GPs prefer private practice or salaried positions? *Scandinavian Journal of Primary Health Care,* 30**,** 229-233.
PEDERSEN, L. B., KJAER, T., KRAGSTRUP, J. & GYRD-HANSEN, D. 2012. General practitioners' preferences for the organisation of primary care: A discrete choice experiment. *Health Policy,* 106**,** 246-256.
SCOTT, A., WITT, J., HUMPHREYS, J., CATHERINE, J., GUYONNE, K., JEON, H. & MCGRAIL, M. 2013. Getting Doctors into the Bush: General Practitioners' Preferences for Rural Location. *Social Science & Medicine,* Forthcoming.

Figure 1

	Practice A	Practice B
Type of practice	Fixed salary	Private practice
Opportunity to control own working hours	Very good	Limited
Opportunity for own professional development	Limited	Very good
Degree of professional autonomy	Limited	High degree
Level of income	150.000 more than your present income	300.000 more than your present income

Which practice do you find most attractive of A and B?

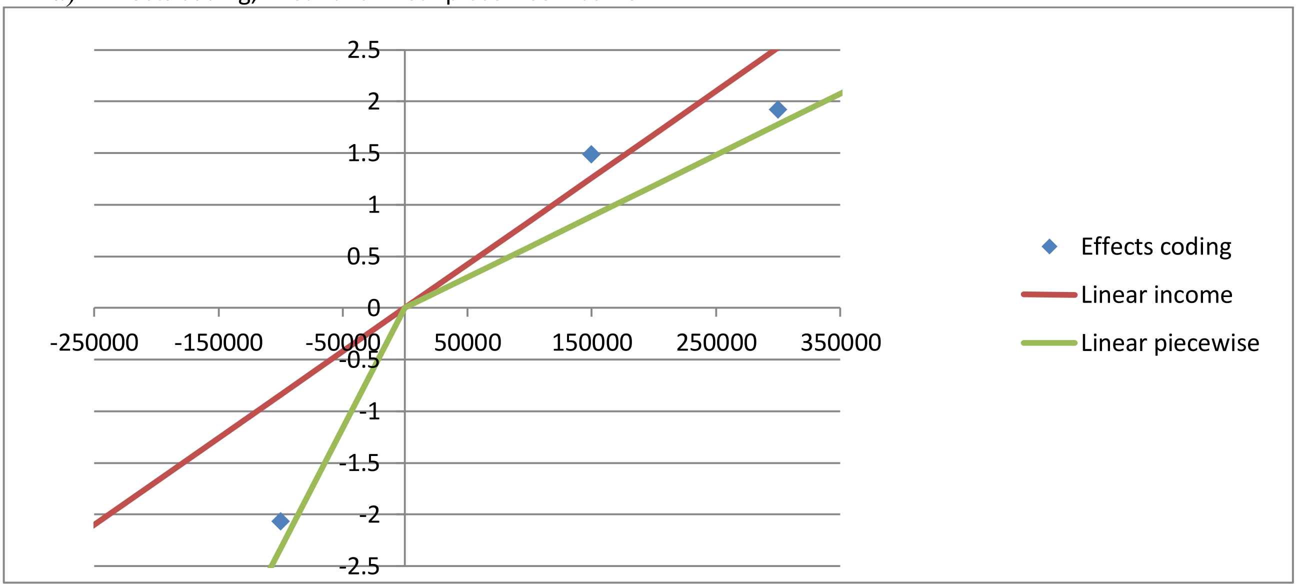
□ Practice A □ Practice B

If you could choose, which practice would you prefer of A, B and your current practice?

□ Practice A □ Practice B □ Your current practice

Figure 2 Changes in utility as a function of changes in income relative to reference category (i.e. current income)

a) Effects coding, linear and linear piecewise income



Appendix 1

Part 3: Your valuation of each individual characteristic WTA format

In this section we want to know how much you value each individual characteristic.

How much more income would be needed to make practice B as attractive as practice A? Please enter necessary increase in income in the box bellow each table

	Practice A	Practice B
Opportunity to control own working hours	Excellent	Limited
Income	Current Income	?

	Practice A	Practice B
Opportunity for own professional development	Excellent	Limited
Income	Current Income	?

	Practice A	Practice B
Degree of professional autonomy	High degree	Limited
Income	Current Income	?

Part 3: Your valuation of each individual characteristic WTP format

In this section we want to know how much you value each characteristics individually.

How much of a reduction in income would you be willing to accept in practice B to make it as attractive as practice A? Please enter the reduction in income you are willing to accept in the box bellow each table.

	Practice A	Practice B
Opportunity to control own working hours	Limited	Excellent
Income	Current Income	?

	Practice A	Practice B
Opportunity for own professional development	Limited	Excellent
Income	Current Income	?

	Practice A	Practice B
Degree of professional autonomy	Limited	Excellent
Income	Current Income	?