

A Comparison of Professional and Private Time Preferences of General Practitioners

Abstract

In a perfect agency relationship, doctors consider all information and select the patient's 'utility maximising' option given the patient's preferences. The patient's time preferences are important as treatments vary in the timing and length of their benefits. However, doctors often do not have full information on patients' preferences and may apply their own preferences. This has generated empirical interest in estimating doctors' time preferences. However, these studies generally elicit doctors' private preferences (preferences for their own health) rather than professional preferences (preferences for the patient). We hypothesise that private and professional preferences may differ. Professional time preferences may be 'taught' in medical school or learned through repeated interactions with patients. If preferences differ then estimates of doctors' private preferences are less informative for medical decision-making. This study compares private and professional time preferences for health in a national sample of General Practitioners, using a between sample design. Time discounting is explored using exponential and quasi-hyperbolic models. We elicit time preferences using multiple price lists. We find no significant difference between the time preference for the self or the patient. This result holds for axiomatic discounting classification and maximum likelihood estimates. We do not find evidence of present-bias. There are a high proportion of increasingly impatient GPs, potentially implying a maximum 'willingness to wait' for treatment benefits. GPs value the health state differently between themselves or for a patient. These results suggest that we can use estimates of private preferences from doctors to inform medical decision-making.

Keywords:

Time preferences; agency; doctors

JEL Codes:

D90, I12

1. Introduction

Since before Arrow (1963), the asymmetry of information between doctors and patients has been recognised in the profession. Understanding the doctor's role as the patient's agent, guiding their choice, is paramount to analysing medical decision-making. In a perfect agency relationship, doctors consider all available alternatives and select the patient's 'utility maximising' option given the patient's preferences. This includes their time and risk preferences. However, the agency relationship is imperfect. Doctors do not have full information on patients' preferences and may apply their own preferences. This has generated empirical interest in estimating doctors' preferences [removed for anonymization] (Gyrd-Hansen, 2002) and in understanding differences in preferences between doctors and their patients (Jahng, Martin, Golin, & DiMatteo, 2005). Recent work by [removed for anonymization] explicitly compares matched patients and doctors, finding patients and their doctors have similar risk preferences, but different time preferences. However, these studies generally elicit doctors' private preferences: preferences for their own health. We hypothesise that doctors have both private preferences and professional preferences (preferences for their patients) and that these may differ. For example, professional preferences may differ because they are 'taught' in medical school and/or learned through repeated interactions with patients. If doctors use professional preferences in the medical decision-making context, then estimates of their private preferences are less informative. This paper directly compares private and professional time preferences in doctors.

We concentrate on time preferences for health as these determine decisions about treatments that vary in the timing and length of their benefits (Fuchs, 1982; Gafni & Torrance, 1984). There is a growing literature on eliciting time preferences but only a few studies have elicited time preferences in doctors (Gyrd-Hansen, 2002; Höjgård et al., 2002; Redelmeier & Heller,

1993). Redelmeier & Heller (1993) investigated doctors' and medical students' time preferences for themselves, using a standard gamble framework to elicit indifference points. A majority (62.1%) of their respondents displayed an exponential discount rate of zero. Within a sample of health students and doctors, Gyrd-Hansen (2002) asked questions from both private and professional perspectives. However, the use of different elicitation methods and health outcomes means the estimated rates are not comparable. Höjgård et al. (2002) elicited professional time preference for health in doctors by obtaining the probability of death during an operation that would make a doctor indifferent to recommending surgery or not to an entire group of patients. This result was used to estimate the implied time discount rate in an exponential model. The 95% confidence interval for GPs' discount rates was 13.58 – 20.46; for 60-year old men it was 11.85 – 20.07. [removed for anonymization] elicited time preferences from doctors in Greece for money and for their own health. Across studies, the estimated range of the exponential discounting parameter is wide, from 0.021 [removed for anonymization] to 0.21 (Højgård et al., 2002).

A few previous studies have examined the impact of perspective (own versus others) on time preferences in general population or student samples. Lazaro et al. (2001) explore the private and social time preferences of law students, finding a statistically significant difference between private and social time preferences for health. However, [removed for anonymization] found similar rates for private and social preferences in the general public. Given the mixed results, and that these results cannot necessarily be generalised to doctors, it is important to directly compare the private and professional time preferences of doctors.

Ours is the also one of first studies to investigate the role of quasi-hyperbolic discounting (Laibson, 1997) in doctors. Quasi-hyperbolic discounting allows for time inconsistency and

has been hypothesised (Elliott, Shinogle, Peele, Bhosle, & Hughes, 2008) and investigated as a factor in health behaviours and outcomes [removed for anonymization](Khwaja, Silverman, & Sloan, 2007). Evidence suggests that many individuals are quasi-hyperbolic discounters (Frederick, Loewenstein, & O'Donoghue, 2002) although this has been debated more recently (Andersen, Harrison, Lau, & Rutström, 2014; Bleichrodt, Gao, & Rohde, 2016). Time consistency is important for medical decision-making as it may explain non-adherence: patients not following through with treatment recommendations that they agreed to. Doctors may recognise that several patients are present biased and may make treatment recommendations in line with the patient's present biased preferences. However, time inconsistency may be undesirable from the perspective of the past self (Bénabou & Pycia, 2002; Fudenberg & Levine, 2006). Doctors may therefore make treatment recommendations that reflect time consistent preferences and encourage patients to do what they planned for their future selves. Only one previous study has examined quasi-hyperbolic discounting in doctors [removed for anonymization]. The estimated present bias parameter in a sample of Greek hospital doctors was 1.04 for health and 0.99 for money suggesting that on average doctors do not exhibit significant present bias.

We compare private and professional time preferences in a national sample of Scottish General Practitioners (GPs). GPs are primary healthcare workers who often act as the gatekeeper to hospital services. They are involved in helping patients manage long-term conditions such as back pain where there are clear trade-offs in the intensity and timing of different treatments. They are therefore more likely to be familiar with the concept of trading costs and benefits over time. Furthermore, GPs' frequent interactions with their chronic patients could make it more efficient for them to maintain separate private and professional preferences.

This paper also examines the impact of the GP's characteristics (such as gender and years since qualifying) and the characteristics of their practice (such as number of patients on the list per GP) on their time preferences for health. This may provide important insights into potential underlying mechanisms for variations in clinical practice.

The remainder of the paper is structured as follows. In Section 2 we review time preference theory, before setting out the analytical methods in Section 3. Section 4 contains the main results. Section 5 then discusses these and Section 6 concludes.

2. Time preferences

Time preferences model how individuals choose between outcomes that occur over time.

Imagine an individual has a preference relation \succsim representing their ordering over pairs of outcomes available at given times, and two outcomes X and Y ($X \prec Y$ if both available the same time) can be obtained at time t and s ($t < s$) respectively. If this person is indifferent

between these two: $(X, t) \sim (Y, s)$, then the preferences exhibit stationarity (Time

Consistency, TC) if the relationship is maintained as both options are shifted forward in time

by an equal amount σ . Stationary preferences can be represented by an exponential discount

function (Prelec, 2004):

$$(X, t) \sim (Y, s) \text{ and } (X, t + \sigma) \sim (Y, s + \sigma) \quad (1)$$

$$D(\delta, t) = \frac{1}{(1+\delta)^t} \quad (2)$$

Where δ is the time preference rate. When δ is low, the discount factor $D(\cdot)$ is high. An

individual who places a lot of weight on future outcomes relative to today has a low δ and is

described as more patient.

Preferences exhibit decreasing impatience (DI) when shifting the options forward means the person now prefers the larger more delayed reward, and this condition is necessary for hyperbolic discounting (Prelec, 2004):

$$(X, t) \sim (Y, s) \text{ and } (X, t + \sigma) < (Y, s + \sigma) \quad (3)$$

Intuitively, this profile of choices can be justified by reasoning such as “what’s one more day?” – the relative weight of the extra length between rewards in comparison to the delay to the first available reward is smaller (Read & Roelofsma, 2003). Lastly, preferences might exhibit increasing impatience (II), where the individual becomes less willing to wait as the outcomes are shifted forward in time:

$$(X, t) \sim (Y, s) \text{ and } (X, t + \sigma) > (Y, s + \sigma) \quad (4)$$

Increasing impatience is less discussed than decreasing impatience and lacks a strong theoretical basis. The non-exponential quasi-hyperbolic form can be used for increasingly or decreasingly impatient individuals (Laibson, 1997):

$$D(\beta, \delta, t) = \begin{cases} 1 & \text{if } t = 0 \\ \beta * \frac{1}{(1+\delta)^t} & \text{if } t > 0 \end{cases} \quad (5)$$

β is a measure of present bias. When β is less than one, the individual exhibits present bias: they care more about very near term rewards than future rewards. However, they weight two points in the further-future consistently (as if an exponential discounter). This reflects decreasing impatience. A β larger than 1 indicates increasing impatience. Events in the near future are valued higher than events today, but two future outcomes are consistently weighted.

3. Methods

We use a split sample design to compare private and professional time preference for health in a sample of Scottish GPs. Data were collected using a postal questionnaire with

personalised cover letter (Appendices 1 & 2). It was not possible to collect the data through online surveys. It is not known what the preferred mode of data collection is in time preference experiments.

Time preferences

Time preferences are elicited using the Multiple Price List (MPL), a standard method to ascertain time preferences outlined in Collier and Williams (1999). We ask the doctor to assume either they, or the patient, are in a given health state. They choose between two treatments, associated with a number of days in full health with certainty, at different points in time. Our health state symptoms are intended to reflect mild back pain: “Mild (chronic) pain that restricts some daily activities, such as your normal work and leisure activities. Otherwise, you have no restrictions in walking about and are not anxious or depressed”. The health state is mild to avoid potentially non-standard behaviour observed around treatments in a severe health state (Chapman, 2005). After treatment, the respondent, or the patient, returns to the original health state. This is aligned to the approach taken in the money domain: by assumption, a prize is spread over a given period before the individual returns to their original income. This design also places us firmly in the domain of gains. We wanted to minimise negative discounting behaviour associated with health losses due to anticipation [removed for anonymization]. The questions are also abstract to reduce the likelihood of respondents telling us that certain treatments do not exist or would not be appropriate, even if they were empirically necessary to identify time preferences.

There are 10 choices within each MPL. We use the same set of payoffs in each block (see Example Block 3) to allow for axiomatic testing. Our early reward is always 14 days in full health. The later rewards start at 14 days as a check for negative discounting ($\delta < 0$). Our

maximum number of future days in full health is 23, implying a weekly discount rate of more than 28% for a 2 week delay and 13% for a 4 week delay if a respondent never switches. The questions are hypothetical as ethically it is not possible to apply treatment preferences in practice. There is mixed evidence on the impact of incentivisation on time preference in the monetary domain (Andersen et al., 2014; Holt & Laury, 2002; Kirby & Maraković, 1996). It should be noted that the interest in this study is the difference between private and professional time preference rather than the mean rate.

[Figure 1 here]

Most studies using the MPL take a structural approach to estimating parameters of a discounting function. By careful construction, blocks of questions can be used to identify axiomatic forms of preference, as in [removed for anonymization] and Halevy (2015). A key advantage of the axiomatic test is that it does not require assumptions about the shape of the utility function for days of full health. Respondents were presented with four blocks (MPLs), allowing two axiomatic tests of stationarity versus present bias (Read & Roelofsma, 2003). We used two ‘delays’ to the early reward (0 or 2 weeks) and two different ‘lengths’ of time between the two rewards (2 or 4 weeks). This gives four blocks (MPLs): Block 1 has 0 delay, length 2 weeks. Block 2 has 0 delay and length 4. Block 3 has delay 2 and length 2. Finally, Block 4 has delay 2 and length 4. This gives two axiomatic tests of stationarity (present bias) by comparing blocks 1 & 3, or 2 & 4 because these pairs have the same lengths and different delays.

The lengths and delays were chosen to reflect clinically relevant decisions made by GPs for minor states such as back pain. In the case of non-specific low back pain, GPs can

recommend fast-acting pharmacological interventions such as ibuprofen or interventions which need longer to take effect, such as group exercise and manual therapies. Similar recommendations exist for other conditions. For osteoarthritis, some treatments that have fast-acting benefits: painkillers; some with delayed benefits: corticosteroid injections; and more-delayed benefits: a managed exercise programme. Our hypothetical treatments therefore reflect real choices facing doctors. We used the back-pain example in the cover letter to explain how GPs trade-off outcomes over time when making treatment decisions (Appendix 1).

To test for differences between private and professional preferences the questions were framed in terms of own or patient health. This type of framing has been successfully used to explore differences in doctors' preferences for themselves and for their patients (Janssen, Teunis, Guitton, & Ring, 2015; Ubel, Angott, & Zikmund-Fisher, 2011). The private arm asked purely 'for you' questions: the respondent was asked to imagine they were in the health state, asked how they would rate the health state on a Visual Analogue Scale (VAS) of 0 (worst health) to 100 (perfect health) for them, and which treatment 'they would prefer' (a point reinforced in each block). In contrast, the professional arm was worded from the point of view of a doctor facing a patient. The respondent was asked to 'imagine a patient' in the health state, how would they rate the health state on a VAS for the patient and which treatment they would recommend to 'this patient', again reinforced in each question block. To check whether the frame worked, the professional frame were asked if they imagined a specific patient. See Appendix 2 for the two surveys.

Other questions

We collected demographic data: year qualified, gender, full- or part-time, practice location (urban or rural), the number of patients on the list and the number of GPs in the practice. We also included risk attitude questions on willingness to engage in financial risks, career and professional risks, and clinical risks based on the MABEL survey (Joyce et al., 2010). The data on risk attitude are used in another paper comparing Scottish GPs and UK qualified GPs in Australia.

Methods of Analysis

We first present the mean switching point across the four blocks by frame. The switching point is defined as the first row the respondent chooses B in a block. With 10 rows, switching row can vary from 1 to 10. We use 11 as a code for individuals that always choose the early days of full health. We conduct axiomatic tests of discounting types by comparing blocks 1 & 3, and 2 & 4. We categorise respondents as follows: Time Consistent (TC) if the switching rows are the same; Decreasingly Impatient (DI) if the individual switches earlier in the later block; Increasingly Impatient (II) if the individual switches later in the later block. We take the responses as strict preferences: if a person switches from A to B in row 6 of block 1 and row 5 of block 3, they are Decreasingly Impatient (DI). This does not allow for any uncertainty or random error in responses. We therefore conduct a sensitivity analysis, widening the definition of TC to ± 1 row. This analysis is exploratory as there is usually a proportion of individuals who cannot be classified as they never switch: they always choose the nearer or later option (in Meier and Sprenger (2015) 40% of respondents never switch). Individuals may appear time consistent because their responses are the same across the compared blocks ('all A' or 'all B'). However, they may have displayed different switching points if a wider range of days of full health in option B had been offered. We also exclude

responses with missing values for the same reason: a single observation in each block does not allow us to observe where the switching point occurred.

We run maximum likelihood (ML) analyses in a similar vein to Andersen et al. (2008; 2014). This approach involves the estimating a utility function's parameters. Initially, linear utility functions were assumed to estimate time preferences. It has been recognised that this may bias the estimates if the utility function is non-linear. Risk preferences are used in the money domain to curve the utility function (Andersen et al., 2008; Laury, McInnes, & Swarthout, 2012). However, due to survey length we were not able to collect a compatible risk preference measure. Linear utility of days in full health may bias estimates of the discount rate upward: it should be noted that the interest lies in the difference between the discount rates across frames, not the rates themselves. Recent evidence from Abdellaoui et al. (2013) suggests there is less non-linearity in the utility function in the time dimension than in the risk dimension. We allow for the non-linearity of the utility function with respect to healthy days in further analysis (details below).

When making intertemporal choices, we model individuals as following one of two discounting approaches: exponential (equation 2) or quasi-hyperbolic (equation 5). Assuming the number of days in full health is a linear reward in the utility function, we treat individuals as if they compare the present values of the two flows, choosing the larger option:

$$PV_j = D(\beta, \delta, t) * X \quad (6)$$

Where j is column A or B, X the size of the reward, β and δ are discounting parameters as above, and t the time at which the reward is available. When β equals one, the discount factor $D(\cdot)$ is $\frac{1}{(1+\delta)^t}$. A more complicated approach would treat the values of the healthy days as separate instantaneous utilities, to calculate a 'true' daily discount rate. While this might be

theoretically more appealing, our primary aim is to compare across samples, not a robust estimate of the doctors' discount rates.

Using our choice data (each row in the survey), we generate an index of the present-value differences and estimate the parameters of the utility function. The index is (Andersen et al., 2014):

$$\nabla PV = \frac{1}{v} (PV_a - PV_b) \quad (7)$$

Where v is a scale parameter. The likelihood function depends on these parameters, the observed choices and personal characteristics (Andersen et al., 2008):

$$\ln L(\beta, \delta, v) = \sum_i ((\ln(\nabla PV) | c_i = 0) + (\ln(1 - \nabla PV) | c_i = 1)) \quad (8)$$

Where i is the individual, c_i the observed choice in a given row (0 indicates a choice in column A, 1 a choice in column B), and responses have standard errors clustered by individual. We test whether $\beta=1$. We explore the effect of framing and demographic variables as linear covariates on the parameters β, δ :

$$\beta_i = \gamma_0 + \sum_K \gamma_k z_k \quad (9)$$

where K is the number of covariates in the model. We have 5 variables other than the frame: gender, full-time, urban practice, thousands of patients per GP and years since qualification. This gives 10 covariates: 5 each on δ and β . We use joint significance tests first on all covariates then on gender and years qualified (proxy for age) to remove covariates that are not jointly significant for more parsimonious models. Gender and age are the two demographic covariates most likely to be associated with time preferences based on previous evidence (Chao, Szrek, Pereira, & Pauly, 2009; Kirby & Maraković, 1996). We test which model of discounting – quasi-hyperbolic or exponential – is a better fit for the data by running two versions of the analysis; one with β varying freely and one without β , constraining it to 1. Goodness of fit is assessed using the Akaike Information Criterion (AIC).

As mentioned above, we do not have a standard measure of risk to separate the effects of risk preference from time preference. We explore the impact of non-linearity of the utility function on β using a Constant Relative Risk Aversion (CRRA) specification:

$$Y(x) = \begin{cases} \frac{x^{(1-\theta)}}{(1-\theta)} & \text{if } t = 0 \\ \frac{\beta}{(1+\delta)^t} \left(\frac{x^{(1-\theta)}}{(1-\theta)} \right) & \text{if } t > 0 \end{cases} \quad (10)$$

Models where θ varies freely do not converge, hence we test three values for θ : 0.1 (Andreoni & Sprenger, 2012; Galizzi et al., 2016), 0.6 (Andersen et al., 2014; Cheung, 2015; Tanaka, Camerer, & Nguyen, 2010), and 1. When θ equals one, CRRA reduces to log-utility and represents strong risk aversion compared to values found in the literature (Ida & Goto, 2009). Without an empirical measure of risk, θ can be interpreted as the decreasing marginal utility of additional healthy days. We also explore the impact of payoff on the estimates by including the VAS score as a multiplicative factor on the number of healthy days: x in (10) becomes $VAS * x$.

The most popular method for correcting for utility curvature uses a risk aversion parameter elicited in a separate task. However, this task is relatively complex and may not always be feasible to include given survey length. This may explain why Andersen et al. (2014) found that only 6 of 20 studies explicitly corrected for curvature in the money domain. When eliciting risk aversion is not feasible, as was the case in our survey, simulation can provide insights into the sensitivity of results to utility curvature. However, the analyses are exploratory only.

Sample

We use a random sample of General Practitioners in Scotland, stratified by health board. A between sample design is used as it is likely that any respondent receiving two very similar questions either would work out the hypotheses, biasing their response, or would fail to appreciate the difference between the surveys and hence replicate responses across frames. Robust sample size calculations for this study are difficult because there are no previous estimates of time preference rates in UK GPs. The closest study is the paper by Anonymous (2016) who elicited time preferences for health using MPLs in a sample of Greek hospital doctors. However, they used a different health outcome and different outcome magnitudes, which are known to influence time preference estimates. These values may, therefore, not be a good approximation of our sample's time preferences. Additionally, there is no clear definition of what a minimally important difference in time preference is. Decisions involving relatively long term outcomes are highly sensitive to the time preference rate: even small differences in time preference rates can lead to different options being chosen. Using a conservative estimate of a 0.0015 (0.15%) difference in time preference rates (δ) and taking the mean and standard error from Anonymous (2016) ($\delta=0.0231$, s.e. 0.0031) the minimum sample size was 138 (69 per group), assuming 80% power.

Response rates are generally low in this population. Assuming a response rate of around 10%, we take a 40% random sample of the 5,000 GPs in Scotland and send out 2,001 surveys to ensure there would be sufficient responses under a conservative return rate. Scotland is split into 14 regional health boards, covering diverse geographical areas. To ensure proportionate coverage in all boards, from the largest (Greater Glasgow and Clyde) to the smallest (Orkney), we stratified on the number of GPs in each board taken from the list of contact details available from Information Services Division (ISD) Scotland. After stratification, each

individual was randomised to receive either the 'private' frame or the 'professional' frame time preferences questionnaire. Data were collected by postal questionnaire. No reminders were sent.

In total, 322 questionnaires were returned (16%): 22 were non-delivered returns, 4 were 'protest' responses and the remaining 296 (159 private and 137 professional) contained at least some response useful for analysis. This useable response rate of 14.8% is higher than anticipated. There was a difference between response rates in each frame (15.9% private and 13.6% professional). This could be because doctors were more willing to make choices for themselves, rather than a hypothetical patient. However, the difference was not statistically significant ($p=0.14$). There were 3 protest responses in the professional frame and one in the private (identified from handwritten notes from respondents). Ethical approval for the study was granted by the College Ethics Review Board, University of Aberdeen and NHS Scotland R&D approval was obtained for all health boards.

4. Results

Respondent characteristics

Table 1 shows descriptive statistics for the sample. The majority of the sample is female, in fulltime employment and based in urban practices. The average respondent left medical school 20 years ago. The demographic variables are not significantly different across frames at the 5% level.

[Table 1 Here]

To explore non-response bias, we compared our sample in terms of gender, urban setting, list size and number of GPs in the practice. The proportion of females, the numbers of patients on their lists, and whether the practice is urban, is similar compared to the sample sent out. Furthermore, the proportion of female responses is not significantly different from the population for all returns, or by frame. This also holds for the urban setting of the practice. However, the average number of GPs per practice is higher for our sample compared to the sample sent out ($p=0.04$). Note that list size was also slightly larger for the sample sent out compared to the full population due to the randomisation procedure. We randomise at health board level, not practice level, and so are more likely to select doctors from practices that have more GPs and patients.

The average VAS score of the health state is higher in the private frame (64.9) compared to the professional frame (58.0). As the data in each frame are not normally distributed (see Figure 2), we used a Kolmogorov-Smirnov test, which rejects equality of the distributions at the 5% level ($D=0.21$, $p=0.01$). The difference in VAS score suggests the framing worked in that GPs were imagining someone other than themselves in the professional frame.

Furthermore, when respondents were asked to describe the patient they imagined only two out of 41 that provided responses noted ‘myself’.

[Figure 2 Here]

Time preferences

Figure 3 shows the proportion of respondents choosing the earlier days of full health as the number of future days of full health increases. In line with other studies, a higher proportion of respondents choose the earlier days of full health if they start now rather than in 2 weeks’

time and if the delay between the earlier and later days of full health is longer. Several respondents never choose the later days of full health, especially for the longer lengths between earlier and later days of full health.

[Figure 3 Here]

Table 2 reports the mean switching row by block. The mean switching row is higher in the private frame suggesting that GPs are more impatient for themselves than for their patients. However, none of the differences in means are statistically significant. Comparing the mean switching row across blocks provides some indication of the impact of delay (starting point) on time preferences. Testing the means across pairs of blocks indicates a difference between private and professional frames. We reject the null hypothesis that the block 1 mean is equal to the block 3 mean in the professional frame ($p=0.03$), but not the private frame ($p=0.40$). However, in block 2 to 4, we reject this hypothesis in both professional ($p=0.02$) and private frames ($p=0.03$). This is an early indication that there could be increasing impatience in the sample.

[Table 2 Here]

Our axiomatic approach classifies the majority of respondents as either time consistent or increasingly impatient. In total, 18.4% of respondents are DI, 34.0% are TC and 47.6% are II. This is similar across the two axiomatic test (2 weeks versus 4 weeks length of time between rewards). The axiomatic test is strict in that it does not allow for randomness or inattention in choices. In part this motivates the ML analysis. We conducted sensitivity analysis by relaxing the definition of TC to be ± 1 row in comparing across blocks. With the relaxed definition,

55.9% are TC with the proportions of DI and II falling to 12.4% and 31.7% respectively. This suggests there is some uncertainty around the axiomatic classification which is addressed in the ML estimation. Note that this excludes individuals missing data (47) or censored in both blocks of a pair. There were 31 individuals with censored observations in the block 1 to 3 comparison, and 70 in the block 2 to 4 comparison. Of these, 75 individuals were censored or had missing data in all four blocks.

Comparing discounting types across professional and private frame, a chi-square test on the distributions for blocks 1 to 3 indicates that the distributions are the same ($X^2=1.08$, d.f.=2, $p=0.58$). This contrasts with the result above that the mean switching row in blocks 1 and 3 is different in the professional frame but not the private frame. However, the results cannot be directly compared as the axiomatic approach excludes a substantial proportion of respondents and different statistical tests are used. The full set of axiomatic results are available in Appendix 3, Table A.1. A weakness of the axiomatic approach is that it excludes individuals that are censored or have missing data. Structural estimation can therefore exploit more of the data.

Table 3 shows the results of the maximum likelihood estimations. The 39 individuals who had censored observations across all four blocks (non-switchers) are included in this analysis. In the model with no demographics $\beta = 1.037$ (95% CI: 1.005, 1.069), $\delta = 0.017$ (95% CI: 0.015, 0.019) in the quasi-hyperbolic model and $\delta = 0.017$ (95% CI: 0.015, 0.018) in the exponential model. There is some evidence of aggregate increasing impatience in the sample. We reject $\beta = 1$ at the 5% level ($p = 0.024$). Framing did not have statistically significant effects on β or δ suggesting that under the assumption of linear utility in days of full health

there is no difference on average between how doctors would choose treatments for themselves or for their patients.

[Table 3 Here]

Similar results are found when the demographics are included. In the full model none of the demographic covariates were individually significant and they were also not jointly significant (Appendix 3 Table A.2). A joint significance test on a model including all 10 covariates did not reject H_0 at the 10% level ($p=0.45$, $d.f.=10$). Taking β and δ separately, we do not reject covariates on δ all being jointly zero ($p=0.27$, $d.f.=5$); nor the β covariates ($p=0.92$, $d.f.=5$). The model therefore includes gender and years qualified only which were jointly significant in both the quasi-hyperbolic model and the exponential model. In both models, being female is statistically significant at the 5% level. Women have a lower δ (more patient), in accordance with previous findings (Kirby & Maraković, 1996; Meier & Sprenger, 2010).

Table 4 shows the maximum likelihood estimates when assuming non-linear utility. The results are qualitatively similar to Table 3: β remains larger than one and can be rejected at the 5% level. The frame remains insignificant in all models. As expected, increasing the curvature of the utility function reduces the point estimates of δ (Model 1: 0.015, Model 3: 0.006) and β (Model 1: 1.034, Model 3: 1.013). This is in keeping with previous results that non-linearity in the utility function biases estimates of time preferences.

[Table 4 Here]

5. Discussion

Our results show that doctors have similar time preferences for themselves and their patients, in terms of rates and underlying discounting types. It may be that our framing was unsuccessful and GPs imagined themselves in the professional frame rather than their patient. In an abstract and cognitively complex task it may be easier for participants to apply their own preferences. However, the clear differences in VAS distributions imply framing was successful. The task was also deliberately abstract: providing more detailed information in the professional frame such as the specific medical condition, age or gender of the patient may influence the time preference estimates due to framing effects. Perspective, own versus professional, was made salient throughout the questionnaire. Further research should explore whether the results are sensitive to different types of framing.

The results showed that doctors perceive a health state as more severe when experienced by a patient rather than themselves. This is important in the medical context as treatment recommendations may vary depending on whether the GP is using professional or private health state values. It is unclear whether their professional or private preferences are better aligned with the patient's health state values and this would be an interesting line of future enquiry.

The paper also contributes to the debate on whether time preferences are appropriately modelled by quasi-hyperbolic or exponential discounting. On average, GPs had a β close to one. There is some evidence that on average individuals are increasingly impatient and the axiomatic test showed that a substantial proportion of the sample is increasingly impatient: 47.6% of doctors as increasingly impatient (future-bias). The proportion of increasing impatience is high compared to most studies in the monetary domain. Meier and Sprenger

(2010) found 9% of their primary sample were future-biased and Halevy (2015) found statistically insignificant evidence of increasing impatience. Augenblick et al. (2015) found 17% of subjects are future biased in money, and 33% are for their effort task. In the health domain, Bleichrodt et al. (2016) found 25% of respondents were increasingly impatient in a sample of 36, using a Time Trade Off method. They found that hyperbolic, rather than quasi-hyperbolic, discounting represented the data better.

While time consistency and decreasing impatience have widely discussed interpretations, increasing impatience does not. Increasing impatience implies that as we shift two outcomes forward in time by equal amounts, one now prefers the smaller sooner option to the larger later one. A potential explanation is that doctors have a maximum ‘willingness to wait’ for treatment benefits to begin because the patient experiences ill health during the initial delay. They become more impatient as outcomes are shifted forwards because they want to move the patient away from a ‘bad state’ by a certain time after the patient first arrives in the clinic, similar to the finding that there are ‘maximal endurable’ times in a health state (Stalmeier & Verheijen, 2013).

Compared to other studies, our mean estimates of δ are in line with health studies using MPLs [removed for anonymization], and comparable to those in the monetary domain (0.021, Tanaka et al. (2010); 0.07, Andersen et al. (2014)). Our estimate of β is relatively high, with the majority of studies finding β either equal to 1 (Andersen et al. (2014), money domain; [removed for anonymization], health domain), or lower (0.6, Tanaka et al. (2010); 0.7, Meier and Sprenger (2015); 0.89, [removed for anonymization]). However, our estimate of 1.037 in the base model (without individual covariates) is remarkably close to [removed for

anonymization], the only other study that elicited present-bias in doctors. Our result adds to the evidence *against* present-bias in doctors making choices over time in the health domain.

There are several limitations to the study. We assume utility linear in healthy days and find that doctors have the same time preferences for themselves and their patients. However, there is a statistical difference in how the sample viewed the same health state applied to themselves or a patient. Whilst the VAS uses a single open-ended question only to elicit health state values, it is a well-established and reliable method for eliciting ratings for health states (Krabbe et al., 2017; McCormack, de L. Horne, & Sheather, 1988; Parkin & Devlin, 2006). Other studies suggest that individuals discount smaller rewards more steeply than larger ones (Andersen, Harrison, Lau, & Rutström, 2013; Benzion, Rapoport, & Yagil, 1989; Kirby & Maraković, 1996). We therefore expect ‘professional’ frame doctors to have lower discount rates (be more patient) than those in the ‘private’ frame because the professional frame perceive the reward as larger. We found no difference in the analysis. Second, as mentioned above similar estimates may be the result of doctors not making decisions from the correct perspective. It may also be that the similarity is caused by mid-point bias: respondents are more likely to switch near the middle (Andersen, Harrison, Lau, & Rutström, 2006). Assuming this effect did not vary by frame because both use the same scale, this would bias our survey against finding a difference between frames.

This conflict demonstrates that applying estimated discount rates for health involves uncertainty for the researcher. In the money domain, this is addressed by using incentivised risk questions to adjust the utility function. In the health domain this is more complex. Fourthly, despite offering a wide range of implied weekly discount rates in the MPLs (up to 28%), a relatively large number of respondents never switched, especially in block 4. For

these individuals, more limited information is available on their time preferences. However, the proportion is similar to other studies (Meier & Sprenger, 2015).

6. Conclusion

This study is the first to directly compare time preferences for the self or a patient in a national sample of General Practitioners, and the first to explore quasi-hyperbolic discounting in GPs. The study used a between sample design to elicit time preferences from doctors for a hypothetical chronic ill-health state that could be cured for given lengths of time. We find a significant difference in how health states are perceived. However, there is no significant difference between the time preference for the self or the patient. We found evidence of increasing impatience in both frames, an understudied aspect of discounting that could be particularly relevant to health outcomes. This study therefore also adds to the evidence against present-bias as a widespread phenomenon. These results suggest that we can use estimates of private preferences from doctors to inform medical decision-making. Future research is clearly required in terms of the impact differences in health state values have on time preferences and in examining whether the results hold across specialties.

References

- Abdellaoui, M., Bleichrodt, H., l'Haridon, O., & Paraschiv, C. (2013). Is there one unifying concept of utility? an experimental comparison of utility under risk and utility over time. *Management Science*, 59(9), 2153-2169.
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, 76(3), 583-618.
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2013). Discounting behaviour and the magnitude effect: Evidence from a field experiment in denmark. *Economica*, 80(320), 670-697.
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2014). Discounting behavior: A reconsideration. *European Economic Review*, 71(0), 15-33.
doi:<http://dx.doi.org/10.1016/j.euroecorev.2014.06.009>
- Andersen, S., Harrison, G., Lau, M., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, 9(4), 383-405. doi:10.1007/s10683-006-7055-6
- Andreoni, J., & Sprenger, C. (2012). Estimating time preferences from convex budgets. *American Economic Review*, 102(7), 3333-56.
- Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*, 53(5), 941-973.

- Augenblick, N., Niederle, M., & Sprenger, C. (2015). Working over time: Dynamic inconsistency in real effort tasks. *The Quarterly Journal of Economics*, *130*(3), 1067-1115. doi:10.1093/qje/qjv020
- Bénabou, R., & Pycia, M. (2002). Dynamic inconsistency and self-control: A planner–doer interpretation. *Economics Letters*, *77*(3), 419-424. doi:[http://dx.doi.org/10.1016/S0165-1765\(02\)00158-1](http://dx.doi.org/10.1016/S0165-1765(02)00158-1)
- Benzion, U., Rapoport, A., & Yagil, J. (1989). Discount rates inferred from decisions: An experimental study. *Management Science*, *35*(3), 270-284.
- Bleichrodt, H., Gao, Y., & Rohde, K. I. (2016). A measurement of decreasing impatience for health and money. *Journal of Risk and Uncertainty*, *52*(3), 213-231.
- Chao, L. W., Szrek, H., Pereira, N. S., & Pauly, M. V. (2009). Time preference and its relationship with age, health, and survival probability. *Judgment and Decision Making*, *4*(1), 1-19.
- Chapman, G. B. (2005). Short-term cost for long-term benefit: Time preference and cancer control. *Health Psychology*, *24*(4S), S41-S48.
- Cheung, S. L. (2015). Risk preferences are not time preferences: On the elicitation of time preference under conditions of risk: Comment. *American Economic Review*, *105*(7), 2242-60.
- Coller, M., & Williams, M. (1999). Eliciting individual discount rates. *Experimental Economics*, *2*(2), 107-127. doi:10.1023/A:1009986005690

- Elliott, R. A., Shinogle, J. A., Peele, P., Bhosle, M., & Hughes, D. A. (2008). Understanding medication compliance and persistence from an economics perspective. *Value in Health*, *11*(4), 600-610.
- Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351-401.
- Fuchs, V. R. (1982). Time preference and health: An exploratory study. *Nber*, WP539
doi:10.3386/w0539
- Fudenberg, D., & Levine, D. K. (2006). A dual-self model of impulse control. *The American Economic Review*, , 1449-1476.
- Gafni, A., & Torrance, G. W. (1984). Risk attitude and time preference in health. *Management Science*, *30*(4), 440-451.
- Gyrd-Hansen, D. (2002). Comparing the results of applying different methods of eliciting time preferences for health. *The European Journal of Health Economics*, *3*(1), 10-16.
- Halevy, Y. (2015). Time consistency: Stationarity and time invariance. *Econometrica*, *83*(1), 335-352.
- Höjgård, S., Enemark, U., Lyttkens, C. H., Lindgren, A., Troëng, T., & Weibull, H. (2002). Discounting and clinical decision making: Physicians, patients, the general public, and the management of asymptomatic abdominal aortic aneurysms. *Health Economics*, *11*(4), 355-370.
- Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, *92*(5), 1644-1655.

- Ida, T., & Goto, R. (2009). Simultaneous measurement of time and risk preferences: Stated preference discrete choice modeling analysis depending on smoking behavior*. *International Economic Review*, 50(4), 1169-1182. doi:10.1111/j.1468-2354.2009.00564.x
- Jahng, K. H., Martin, L. R., Golin, C. E., & DiMatteo, M. R. (2005). Preferences for medical collaboration: Patient–physician congruence and patient outcomes. *Patient Education and Counseling*, 57(3), 308-314.
- Janssen, S. J., Teunis, T., Guitton, T. G., & Ring, D. (2015). Do surgeons treat their patients like they would treat themselves? *Clinical Orthopaedics and Related Research®*, 473(11), 3564-3572.
- Joyce, C. M., Scott, A., Jeon, S., Humphreys, J., Kalb, G., Witt, J., & Leahy, A. (2010). The "medicine in australia: Balancing employment and life (MABEL)" longitudinal survey - protocol and baseline data for a prospective cohort study of australian doctors' workforce participation. *BMC Health Services Research*, 10(1), 1-10. doi:10.1186/1472-6963-10-50
- Khwaja, A., Silverman, D., & Sloan, F. (2007). Time preference, time discounting, and smoking decisions. *Journal of Health Economics*, 26(5), 927-949.
- Kirby, K. N., & Maraković, N. N. (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic Bulletin & Review*, 3(1), 100-104.
- Krabbe, P. F., Stolk, E. A., Devlin, N. J., Xie, F., Quik, E. H., & Pickard, A. S. (2017). Head-to-head comparison of health-state values derived by a probabilistic choice model and

- scores on a visual analogue scale. *The European Journal of Health Economics*, 18(8), 967-977.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2, In Memory of Amos Tversky (1937-1996)), 443-477.
- Laury, S. K., McInnes, M. M., & Swarthout, J. T. (2012). Avoiding the curves: Direct elicitation of time preferences. *Journal of Risk and Uncertainty*, 44(3), 181-217.
- Lazaro, A., Barberan, R., & Rubio, E. (2001). Private and social time preferences for health and money: An empirical estimation. *Health Economics*, 10(4), 351-356.
- McCormack, H. M., de L. Horne, D. J., & Sheather, S. (1988). Clinical applications of visual analogue scales: A critical review. *Psychological Medicine*, 18(4), 1007-1019.
doi:10.1017/S0033291700009934
- Meier, S., & Sprenger, C. D. (2015). Temporal stability of time preferences. *Review of Economics and Statistics*, 97(2), 273-286.
- Meier, S., & Sprenger, C. (2010). Present-biased preferences and credit card borrowing. *American Economic Journal: Applied Economics*, 2(1), 193-210.
- Parkin, D., & Devlin, N. (2006). Is there a case for using visual analogue scale valuations in cost-utility analysis? *Health Economics*, 15(7), 653-664.
- Prelec, D. (2004). Decreasing impatience: A criterion for non-stationary time preference and “Hyperbolic” discounting. *The Scandinavian Journal of Economics*, 106(3), 511-532.

Read, D., & Roelofsma, P. H. M. P. (2003). Subadditive versus hyperbolic discounting: A comparison of choice and matching. *Organizational Behavior and Human Decision Processes*, 91(2), 140-153. doi:[http://dx.doi.org/10.1016/S0749-5978\(03\)00060-8](http://dx.doi.org/10.1016/S0749-5978(03)00060-8)

Redelmeier, D. A., & Heller, D. N. (1993). Time preference in medical decision making and cost-effectiveness analysis. *Medical Decision Making : An International Journal of the Society for Medical Decision Making*, 13(3), 212-217.

Stalmeier, P. F. M., & Verheijen, A. L. (2013). Maximal endurable time states and the standard gamble: More preference reversals. *The European Journal of Health Economics*, 14(6), 971-977. doi:10.1007/s10198-012-0445-0

Tanaka, T., Camerer, C. F., & Nguyen, Q. (2010). Risk and time preferences: Linking experimental and household survey data from vietnam. *American Economic Review*, 100(1), 557-571.

Ubel, P. A., Angott, A. M., & Zikmund-Fisher, B. J. (2011). Physicians recommend different treatments for patients than they would choose for themselves. *Archives of Internal Medicine*, 171(7), 630-634.