THE UNIVERSITY OF MELBOURNE

FACULTY OF BUSINESS & ECONOMICS

Melbourne Institute Working Paper Series

Working Paper No. 5/17

Motivation and Competition in Health Care

Anthony Scott and Peter Sivey

MELBOURNE INSTITUTE®
of Applied Economic and Social Research

Motivation and Competition in Health Care*

Anthony Scott[†] and Peter Sivey[‡]
[†] Melbourne Institute of Applied Economic and Social Research,
The University of Melbourne
[‡] School of Economics, Finance and Marketing, RMIT University

Melbourne Institute Working Paper No. 5/17

ISSN 1447-5863 (Online)

ISBN 978-0-73-405240-7

January 2017

* Funding is acknowledged from an ARC Future Fellowship (FT0991296) and an NHMRC Principal Research Fellowship (628471). Funding for MABEL comes from the National Health and Medical Research Council (Health Services Research Grant: 2008-2011; and Centre for Research Excellence in Medical Workforce Dynamics: 2012-2016) with additional support from the Department of Health (in 2008) and Health Workforce Australia (in 2013). The work on competition and prices was supported by an ARC Discovery grant (DP0771005). We thank the doctors who gave their valuable time to participate in MABEL, and the other members of the MABEL research team. The MABEL survey study was approved by the University of Melbourne Faculty of Economics and Commerce Human Ethics Advisory Group (Ref. 0709559) and the Monash University Standing Committee on Ethics in Research Involving Humans (Ref. CF07/1102 - 2007000291). The views in this paper are those of the authors alone. For correspondence, email <a.scott@unimelb.edu.au>.

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* (03) 8344 2100**
***Fax* (03) 8344 2111**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

Abstract

Non-pecuniary sources of motivation are a strong feature of the health care sector and the impact of competitive incentives may be lower where pecuniary motivation is low. We test this hypothesis by measuring the marginal utility of income of physicians from a stated-choice experiment, and examining whether this measure influences the response of physicians to changes in competition on prices charged. We find that physicians exploit a lack of competition with higher prices only if they have a high marginal utility of income.

JEL classification: I11

Keywords: Doctors, incentives, competition, motivation

1. Introduction

Evidence on the impact of competition in health care is mixed, with most research focusing on hospital competition and little research on the markets for physicians (Gaynor and Town 2012). The effect of competition and other economic incentives will depend, in part, upon the objectives of health care providers. The aim of this paper is to examine if the impact of competition is influenced by the motivations of health care providers. In health care, it has long been recognised that these objectives may include sources of motivation other than profit (Arrow 1963). At the heart of many studies of the effects of financial incentives on health care providers' behaviour lies a utility function that includes not only net income (profit) but also patient's health status or utility (Evans 1984, Feldstein 1970, McGuire 2000, Ellis and McGuire 1990, Siciliani 2009, Arrow 1963). Physicians' preferences for net income and patients' health status, and how they are traded-off, are a key source of variation in their responses to financial incentives, and motivated much of the debate in the 1970s and 1980's on supplier-induced demand (Labelle, Stoddart, and Rice 1994).

A more recent parallel literature has developed within behavioural economics that has recognised the importance of sources of motivation other than income. Pro-social motivation, intrinsic motivation, and organisational goals as missions have all been used to recognise that other, non-pecuniary, objectives matter in explaining economic behaviour (Frey 1997, Frey and Jegen 2001, Fehr and Falk 2002, Fehr and Camerer 2007, Besley and Ghatak 2005, Gregg et al. 2011). In addition, in the field of neuroeconomics, the prefrontal cortex in the brain has been shown to play an important role in resolving conflicts and trade-offs between selfish and pro-social rewards (Fehr and Camerer 2007).

If these other non-pecuniary sources of motivation are important, and to the extent that these motives are more prevalent amongst health care professionals, this may reduce the impact of economic incentives in health care markets (Brekke, Siciliani, and Straume 2011, Brekke et al. 2014). Financial incentives may not work for physicians with a relatively low marginal utility of income but a high concern for patients' health or other non-pecuniary factors, or the financial incentive may need to be higher to elicit a behavioural response. In general, if a physician places a relatively high weight to improving patients' health, then there is a need for much less complex remuneration and payment schedules especially if these objectives match the mission of the hospital or physician group practice (Mooney and Ryan 1993, Besley and Ghatak 2005).

However, physician remuneration schedules are becoming more complex over recent years with the introduction of complex pay for performance schemes, suggesting that third-party payers think that physicians are motivated primarily by money. Empirical evidence is, however, mixed. The degree to which physicians react to financial incentives has been a focus of empirical research for decades (McGuire 2000). There is a large literature that examines the effects of changes in remuneration schemes (Gosden et al. 2001), the effect of changes in the level of fees (Clemens and Gottlieb 2014), the effect of changes in the level of wages using labour supply models (Baltagi, Bratberg, and Holmas 2005), and the effect of performance-related pay on physician behaviour (Scott et al. 2011). There is considerable variation in the reported effects of financial incentives on a range of behaviours. If the size of the effect of incentives varies with context, this is consistent with the existence of other unobserved non-pecuniary sources of motivation.

Though the source of motivation has been recognised in theoretical models of health care provider behaviour for decades (Ellis and McGuire 1986), empirical work on the direct measurement of the different sources of motivation for health care providers is still in its infancy. The altruism of medical students and other health workers in training has been measured recently in a number of laboratory experiments (Godager and Wiesen 2013, Hennig-Schmidt, Selten, and Wiesen 2011, Kolstad and Lindkvist 2013, Smith et al. 2012). Some of these have measured altruism and pro-social preferences in the lab (Godager and Wiesen 2013, Hennig-Schmidt, Selten, and Wiesen 2011, Smith et al. 2012). Kolstad and Lindkvist (2013) examined associations between these lab-based measures with career intentions such as public or private sector work. Two studies have examined preferences from lab experiments and their impact on actual behaviour after the experiment. Serra, Serneels, and Barr (2011) found health professionals with a relatively high pro-social motivation were more likely to work in the non-profit sector three years later, and Lagarde and Blaauw (2014) found pro-social nurses were more likely to choose to work in rural areas.

Other studies have focused on heterogeneity in the monetary motivation of physicians. For example, Rizzo and Zeckhauser (2003) and Rizzo and Zeckhauser (2007) used a question in a survey that asked doctors what they think their income should be given their career stage, and define this as a reference income (or target income). They found that those whose actual income is below their reference income have stronger growth in income over time, and also show that females do not respond to reference incomes. Iversen and Lurås (2000) used information from the introduction of a capitation scheme in Norway, where physicians had to state their preferred number of patients on their list. Physicians who were allocated less than their preferred number of patients (i.e. a shortage of patients) were assumed to be more highly income motivated and were found to provide more services to their patients.

The growing literature on discrete choice experiments is also a source of evidence on the contents of physician's utility functions and heterogeneity in their preferences for income. These are experiments conducted in postal and online surveys rather than the laboratory where respondents make choices between sets of alternative types of job. Job characteristics are varied according to an experimental design and an indirect utility function is estimated that contains the earnings of health professionals and a range and other job characteristics (Lagarde and Blaauw 2009, Lancsar and Louviere 2008). Discrete Choice Experiments (DCEs) provide direct estimates of the marginal utility of income. A number of studies of physicians have found that the marginal utility of income was statistically significant (Chomitz et al. 1998, Hanson and Jack 2010, Hole and Kolstad 2010, Kolstad 2011, Wordsworth et al. 2004, Scott 2001, Ubach et al. 2003, Gosden, Bowler, and Sutton 2000, Vujicic et al. 2010, Kruk et al. 2010, Scott et al. 2013). Several of these examined heterogeneity in the marginal utility of income due to observable physician characteristics (Chomitz et al. 1998, Scott 2001, Ubach et al. 2003, Wordsworth et al. 2004, Hanson and Jack 2010, Kruk et al. 2010, Hole and Kolstad 2010). These studies found only a few physician characteristics associated with the marginal utility of income, including income, age, experience, and gender. These effects are not consistent across studies, though several studies found men to have a higher marginal utility of income than women physicians.

The aim of this paper is to add to these literatures by examining whether the degree of monetary motivation modifies the impact of economic incentives on physicians. This has not yet been examined empirically in health care. In particular, we examine the pricing incentives given by the level of competition between physicians. Physicians with a low monetary motivation will place a lower weight on any changes in revenue and profit caused by changes in competition. They are therefore less likely to respond to competition by lowering their prices compared to physicians with a high monetary motivation. We find empirical results which support this hypothesis.

We examine these issues in the Australian health care system which is a useful context to study competition between physicians. For general practitioners (GPs) there are no restrictions on patient choice of GP, no restrictions on physician mobility (unless a GP enters from overseas), and no price regulation in the fee for service system so GPs can charge what the market will bear. Our data come from the unique Medicine in Australia: Balancing Employment and Life (MABEL) panel survey of doctors which was used in a previous study examining the impact of competition on the prices charged by GPs in Australia (Gravelle et al. 2016). This paper used an individual measure of competition: the distance between GP practices. A key issue with previous research which has used area-based measures on competition based on doctor to population ratios or market share indices, is that unobserved area characteristics are correlated with both location decisions (competition) and prices. By using an individual measure of competition, Gravelle et al (2016) used area fixed effects to control for unobservable area characteristics. They found that GPs in areas with less competition (longer distances between GPs) charged higher prices.

We extend the models from Gravelle et al (2016) to test whether the effect of competition is influenced by GPs' monetary motivation. Our measure of monetary motivation is a physician-specific measure of the marginal utility of income estimated from a random-parameter mixed logit model using data from a DCE administered to GPs in 2008 (Scott et al. 2013). The DCE was in the context of choosing a job in which to work, which included an attribute of the percentage change in annual earnings, along with other practice characteristics. A DCE has several advantages over traditional hedonic wage equations that use revealed preference data to examine compensating differentials for jobs. In these studies endogeneity bias of job characteristics is a serious issue, both in terms of unobservable job characteristics and in terms of selection on unobservable individual characteristics (Ekeland, Heckman, and Nesheim 2002). In DCEs the impact of unobservable job characteristics are minimized by the researcher having complete control over which job characteristics are included in the experiment. One can include the most important and relevant job characteristics that have been identified from theory, previous research or pilot studies, and which usually do not appear in revealed preference datasets. This ensures that the most salient job characteristics are included.

A second source of endogeneity in hedonic models is due to selection on individual unobservable characteristics. This is not present in a DCE since individual characteristics are orthogonal to the variation of the attribute levels in the experiment because of the experimental design. The attributes are presented exogenously and do not reflect previous choices of current job characteristics as would be the case using revealed preference data. Furthermore, unobservable individual characteristics are allowed to influence preferences by using a mixed logit modelling framework which includes preference heterogeneity through the inclusion of random parameters for each attribute.

As information was collected on each GP's annual earnings, the DCE was used to estimate the marginal utility of an extra dollar. The ability to estimate the marginal utility of income for each individual GP is made possible by the estimation of a mixed logit model, the generalised multinomial logit (GMNL), after which it is possible to use a Bayesian procedure to produce an estimate of the marginal utility of income for each GP based on their responses to the DCE. The marginal utility of income for each GP is then matched to the GPs included in the competition model, and interacted with the measure of competition. The results show that GPs who are faced with a low level of competition will raise their prices more if they have a high level of monetary motivation. For low levels of monetary motivation, the effect of competition on prices becomes statistically insignificant.

The next section outlines in more detail the institutional context of GPs within the Australian health care system. We then briefly describe the data and the two studies on which this paper is based. We then discuss the estimation of the marginal utility of income, and present the results showing how the marginal utility of income influences GPs responses to competition.

2. Institutional context

Australia has a tax-financed universal health care system, Medicare, which provides subsidies to patients for private medical services, including GP services. GPs are organised in small practices, usually in partnerships or owned by companies and are gatekeepers as non-emergency visits to medical specialists or hospitals require a referral from a GP. GPs are paid by patients through fee-for-service, with the patient receiving a fixed subsidy from the Medicare Benefits Schedule (Department of Health and Ageing 2007). GPs can charge what the market will bear – there are no price controls or upper limits on prices charged. This means that there is a varying co-payment - the difference between the price charged and the Medicare subsidy. This gap or out of pocket payment cannot be covered by private health insurance. GPs can choose to charge patients a price equal to the Medicare subsidy so the patient does not pay a co-payment which is known as 'bulk-billing' and around 80% of GP visits are bulk-billed[1]. For patients who aren't bulk-billed, GPs can price discriminate and are able to charge different patients different prices. GPs use this discretion to either bulk-bill patients or charge them a price greater than the subsidy. Usually, prices posted by the practice are charged to all patients as this is administratively easier.

There are four types of subsidy available for the vast majority of GP visits, Level A-D, which vary by complexity and time. Over 80% of visits claimed are for Level B consultations[2] with a Medicare subsidy of around AUD $33 in 2008. The subsidy is slightly higher for children under 16 and for those with concession cards, and there are also a number of safety nets in place which provide higher subsidies for those with very high annual expenditures. In addition to revenue from fee-for-service, GPs receive a range of other payments from Medicare (e.g. for being located in a rural area) that are around 10% of total revenue. There are no entry restrictions (unless the GP is a recent migrant) into geographic areas, and patients can visit any GP as there is no registration or enrolment.

[1] http://www.health.gov.au/medicarestats [accessed 07/10/2016]
[2] http://medicarestatistics.humanservices.gov.au/statistics/mbs_group.jsp [accessed 07/10/2016]

3. Hypotheses

Gravelle et al (2016) develop a model of price competition with vertical and horizontal differentiation to inform the estimations of their empirical models on the same Australian dataset we study in this paper. Their model predicts that markets with higher distances between GP practices will lead to higher prices. Several studies have considered the theoretical implications of variation in the motivation of physicians or healthcare providers more generally. Ellis and McGuire (1986) provide a benchmark principal-agent model where physician's trade-off patient's health benefit with volume of care. Typically, patients' health benefit is maximised at some optimum volume of care and physicians' reimbursement incentives lead to over-treatment or under-treatment relative to this optimum. A higher level of altruism would result in a reduction in volume in the over-treatment case or an increase in volume in the under-treatment case. This model is used by Godager and Weisen (2013) to estimate altruism levels from experimental data on medical students' treatment choices.

In our context pricing rather than volume is the outcome of interest. A similar trade-off exists in that patients will always be better off with lower prices whereas doctors will prefer higher prices conditional on their elasticity of demand being less than unity. The decision to increase prices therefore reflects the trade-off between health benefit and profit suggested by Ellis and McGuire (1986). We can infer that an increase in doctors' monetary motivation, can only lead to no change, or an increase in prices, ceteris paribus. Our first hypothesis to test is that GPs with higher monetary motivation (a higher marginal utility of income) will charge higher prices.

In addition to the direct effect, monetary motivation may interact with the marginal effect of financial incentives on behaviour. Siciliani (2009) includes motivation-crowding in a model where strong financial incentives can themselves reduce the level of altruism (as in the literature surveyed by Frey and Jegen, 2001) and shows how altruism can increase or decrease the effect of financial incentives on the volume of care provided. In our context, the financial incentive is given by the level of local competition faced by each GP, measured by distance to competing practices. GPs facing, for example, a fall in local competition, may increase prices more if they have a higher marginal utility of income than if they have a low marginal utility of income.

Our second hypothesis is that GPs will respond to a fall in local competition by increasing prices more if they have high monetary motivation than if they have lower monetary motivation.

4. Data

We use data from the first wave of the 'Medicine in Australia: Balancing Employment of Life' (MABEL) longitudinal survey of doctors. The methods of the survey are described in more detail elsewhere (Joyce et al. 2010). The GP survey was sent to the population of 22,127 GPs in clinical practice in Australia in 2008 (Wave 1 of MABEL). The response rate for GPs was 17.65% (3,873/22,137) after three reminders. Respondents were broadly representative to the population of GPs in Australia with respect to age, gender, geographic location, and hours worked (Joyce et al. 2010). Our estimation sample of 1698 is 10.3% of the population of 16,382 GPs in urban areas in Australia in 2008 (population data from the Australian Institute of Health and Welfare 2010). The analysis uses data only on GPs in metropolitan areas of Australia as there are extra payments and incentives for doctors in rural areas which can hinder simple interpretation of the market environment. Our estimation sample averages are 48% female (population 40%) age 50.1 years (population 50.5), total hours worked 38.3 (population 37.7). Our sample therefore seems representative with respect to these key variables, with the exception of gender, where female GPs are over-represented.

The estimation sample is also linked to local-area characteristics of the practice of each responding GP using postcodes or Statistical Local Area codes. The 1698 GPs in the estimation sample are located in 605 postcode areas with an average population of 18,588. We use postcode area level data from the 2006 census on the population age distribution, ethnicity, self-reported disability, and socio-economic status measured by the Socio-Economic Index for Areas (SEIFA). The SEIFA Index of Relative Socio-Economic Advantage and Disadvantage is constructed by the Australian Bureau of Statistics from 22 variables measuring education, income, occupational structure, employment status, and family structure. Higher values correspond to greater advantage and we expect postcodes with a higher SEIFA score to have greater valuation of quality and thus to have GPs who set higher prices and provide higher quality.

The GPs in the estimation sample are located in 397 SLAs with an average population of 33,359. We attribute SLA level data on median house prices and population density to GPs via their practice address. House prices may capture higher premises costs for GPs and richer populations who have a higher willingness to pay for GP services. In some SLAs there are additional incentives for bulk-billing and we include a dummy variable to indicate these SLAs.

5. Methods

We adopt a two stage approach:

1. First we estimate a model of GPs income and other workplace preferences using data from a DCE. The empirical model is used to produce estimates of monetary motivation (the marginal utility of income) for each GP and examine its association with observable GP characteristics.
2. Second we incorporate our estimates of GPs' monetary motivation into the empirical model of competition and pricing decisions

The first stage follows the analysis of Scott et al. (2013), who describe in detail the methods and results of the DCE which was included in Wave 1 of the MABEL survey in 2008 and completed by 3,685 GPs. DCEs estimate the parameters of indirect utility functions and the marginal utility of each argument in the utility function (Lancsar and Louviere 2008, Louviere and Lancsar 2009). The DCE methodology is based on random utility theory, the discrete choice analog of utility theory (Manski 1977). If earnings or income is included as an attribute, then the empirical model provides a direct estimate of the marginal utility of income, as well as all other attributes. The utility U_{nij} and choice outcome Y_{nij} of physician n for alternative i from choice set j is:

$$\begin{gathered} U_{nij} = X_{nij}\beta + \varepsilon_{nij} \\ Y_{nij} = 1 \; if \; U_{nij} > U_{n-ij}, \\ or = 0 \; otherwise \end{gathered} \tag{1}$$

$$n = 1,...N; \; i = 1,2; \; j = 1,...J.$$

where X_{nij} is a k-vector of observed attributes of alternative i including income, β is a k-vector of marginal utilities of the attributes, including the marginal utility of income, and ε_{nij} is i.i.d. type 1 extreme value giving the multinomial logit model.

The model includes eight attributes, four with three levels, and four with four levels (Figure 1). The income attribute in the experiment was defined as the percentage change in earnings with levels of -15%, no change and +15%. We modified this attribute by multiplying the three values of the independent variable (-15%, 0, +15%) by each GPs current annual earnings as reported in the survey so the independent variables were expressed in terms of changes in absolute earnings (in $000s) rather than percentages.

The levels of each attribute are varied across the alternatives according to an efficient fractional factorial experimental design. This design produced thirty six alternatives. These were blocked into four groups of nine choices, each with two alternatives. Each GP was randomly allocated to be presented with one of the four versions, each containing nine pairs of alternatives.

Two job alternatives were presented to GPs (A and B) and they were asked which job (A or B) do they prefer (forced choice), and then asked which job they would choose: A, B, and their current job (status quo). The latter 'status quo' option was included to account for potential status quo bias and reference dependent preferences. In practice, GPs chose their current job in 84% of choices (equivalent to 64% of GPs choosing their current job for all nine choices). For scenarios where GPs choose the status quo alternative, no information is revealed about preferences between the two alternative jobs in the experimental design. For this paper it is important to maximise the amount of information about the preferences of each GP to obtain accurate estimates of the marginal utility of income. Therefore, in contrast to Scott et al (2013), for the main specifications we use data from the first forced choice to provide the maximum information about preferences for each GP. We conduct a robustness check on the results using the data including the status-quo option.

Unobserved heterogeneity in marginal utilities can be modelled using an extension of the multinomial logit, the mixed logit model (MIXL):

$$U_{nij} = X_{nij}\beta_n + \varepsilon_{nij} \qquad (2)$$

$$\beta_n = \tilde{\beta} + \eta_n$$

where η_n is a k-vector of mean-zero individual -specific deviations from the mean marginal utility such that β_n is a k-vector of individual-specific marginal utilities of each attribute with a distribution $F(\beta_n;\theta)$ specified by the researcher (Train 2009). The vector of parameters θ (typically the means and standard deviations of the random coefficients β_n) characterises the distribution of β_n.

We estimate an extension of the mixed logit model, the generalised multinomial logit model (G-MNL), (Fiebig et al. 2010) which accounts for the possibility of scale heterogeneity: that the variance of the error terms varies across individuals. The model is extended by modifying β_n:

$$\beta_n = \sigma_n\tilde{\beta} + \gamma\eta_n + (1-\gamma)\sigma_n\gamma\eta_n \qquad (3)$$

where $\sigma_n = exp(\bar{\sigma} + \tau v_n)$, $v_n \sim N(0,1)$, and $\bar{\sigma} = \tau^2/2$, so there are two extra parameters, τ and γ, to be estimated. Scale heterogeneity may exist due to near-lexicographic preferences where marginal utilities for some attributes are very high (ie scaled up); or at the other extreme can be due to randomness of behaviour where the idiosyncratic error term dominates and an individual is very unsure of their choices. Fiebig et al (2010) argue that the G-MNL model is flexible enough to model data from these "extreme" respondents, therefore providing a much better fit to the data.

For the purpose of this study, the important outputs from the GMNL are the individual-level coefficient estimates which are available post-estimation. Where the β_n are distributed according to the distribution function $F(\beta_n; \theta, \tau, \gamma)$ the individual-level coefficients are the expected values of the β_n given the parameter estimates and the choices made by each individual: $E[\beta_n | Y_n, X_n;\ \hat{\theta}, \hat{\tau}, \hat{\gamma})]$ (Greene 2007). The intuition of these estimates is that the preferences of GPs who face nine choice sets X_n will vary according to the specific sequence of alternatives Y_n they choose such that there is a distribution of coefficients β_n. Two GPs who completed the same set of nine choices (X_n), and choose the same alternatives (Y_n), will have the same individual-specific coefficient estimate of β_n (Train 2009). The estimates of β_n can be expressed as follows:

$$E[\beta_n|Y_{n,}X_n;\ \hat{\theta},\hat{\tau},\hat{\gamma})] = \int \frac{\beta \Pr(Y_n|x_{n,}\beta)f(\beta_n;\hat{\theta},\hat{\tau},\hat{\gamma})d\beta}{\Pr(Y_n|x_{n,}\beta)f(\beta_n;\hat{\theta},\hat{\tau},\hat{\gamma}d\beta} \tag{4}$$

where the integral can be approximated by taking draws β^r from the distribution $f(\beta_n;\hat{\theta},\hat{\tau},\hat{\gamma})$, calculating the predicted probabilities $\Pr(Y_n|x_n,\beta^r)$ and generating the simulated $E[\beta_n|Y_{n,}X_n;\ \hat{\theta},\hat{\tau},\hat{\gamma})]$:

$$\check{\beta}_n = \sum_r w^r \beta^r, \tag{5}$$

where the weights w_r are given by:

$$w_r = \frac{\Pr(Y_n|x_n,\beta^r)}{\sum_r \Pr(Y_n|x_n,\beta^r)}$$

We specify a log-normal distribution for the income coefficient to embed an assumption of monotonicity. All other random coefficients are assumed to have a normal distribution in the initial model. A second, more parsimonious model was estimated using fixed coefficients for those coefficients that did not have statistically significant standard deviations in the initial model. This more parsimonious model was then used to estimate the individual-specific income coefficients, or marginal utilities of income. The estimates of individual-specific marginal utility were then standardised to have zero mean and standard deviation of 1 to aid interpretation.

The second stage of the analysis examines the association between the individual GP-specific marginal utility of income estimated in the first stage and the response of GPs' pricing decisions to competition following the method of Gravelle et al (2016). The authors used pricing data from the same survey used to conduct the DCE, Wave 1 (2008) of the MABEL survey. The survey asked questions about prices charged for a 'standard' Level B consultation (which comprise over 80% of all consultations). Three main measures of prices were used, i) the proportion of patients who are bulk-billed F^b, (charged only the level of the Medicare rebate m and face a zero copayment); ii) the average gross price $\overline{p}^{nb} + m$ which is the Medicare rebate m plus the average price paid by patients who are not bulk-billed, and; iii) the average gross price for all patients $m + \left(1 - F^{b}\right)\overline{p}^{nb}$.

To measure competition, Gravelle et al (2016) calculated straight line road distances between each GP survey respondent and the other nearby practices in the population. In the analysis presented here we use the 3rd nearest GP practice, but using the nearest or the 5th nearest does not change the results qualitatively. A key issue in the previous literature is that competition is usually measured at a small area level but where small area characteristics (e.g. schools, availability of hospitals, amenities, practice costs) also influence the number of GPs in that area and therefore competition. Our identification strategy uses distance to the other nearby GP practices which varies both between and within areas, and use area fixed effects to control for all unobserved area-specific factors that influence GPs supply in small areas. The impact of competition is therefore identified from within area differences in distance between GP practices, whilst controlling for all unobserved characteristics of the area. The areas are small enough that unobserved factors varying within areas and which correlated with within area variation in prices and competition are negligible. More detail of the modelling are in Gravelle et al (2016). The results of this earlier study show that GPs with more distant competitors charge higher prices and are less likely to bulk bill (charge zero co-payment).

We estimate the same linear models as in Gravelle et al, allowing the marginal utility of income estimates (estimated $E\left[\beta_n | Y_{n,} X_n;\ \hat{\theta}, \hat{\tau})\right]$ for the income attribute which we rename MUY) to enter as an explanatory variable to test whether it has a direct effect on competition. We also include it as an interaction term with the measure of competition to test our main hypothesis that the effect competition is greater for physicians with a high monetary motivation. Extending the baseline model (equation 13) from Gravelle et al (2016):

$$y_{nr} = \beta_0 + \beta_1 GPdist_{nr} + \beta_{11} GPdist_{nr} * MUY_{nr} + \beta_{12} MUY_{nr} + \beta_2 GPchars_{nr} + \beta_3 Areachars_r + \epsilon_{nr} \tag{6}$$

The dependent variable y_{nr} is one of the three alternative measures of prices. $GPdist_{nr}$ is the distance to the third-nearest other GP practice in the population and it is interacted with MUY_{nr}, the estimated individual GP-level income coefficient from the first stage. The models also include a set of GP characteristics, $GPchars_{nr}$ (age, gender, spouse, dependent children, Australian medical graduate, years of experience, whether a partner in the practice (self-employed), and whether the practice is taxed as a company). We also include the characteristics of the local geographic area (Statistical Local Area), $Areachars_{r}$ including an index of socio-economic advantage and disadvantage, median house prices, the proportion of population under 15 and over 65, the proportion disabled, the proportion of immigrants, and population density. For each price variable, we estimate an OLS model, random and fixed area effects models, and a Mundlak specification where area-means of GP-level variables are included as control variables.

If the marginal utility of income is a factor that influences GPs selection into geographic areas, such as areas where expected profits are high due to high demand and/or few GPs, then this will be accounted for through the area fixed effects which capture all unobserved area characteristics. Identification now relies on between and within area differences in both the marginal utility of income and distance to the 3rd nearest GP. To the extent that the marginal utility of income is associated with life cycle and family factors of the GP, these are included in the set of independent control variables in the second stage regression models.

As the estimated marginal utility of income is a generated regressor, the standard errors in the second stage of the analysis are not consistently estimated using standard methods. To address this potential error in our inference, we present robustness checks of the model where we bootstrap the standard errors in the second stage regression by bootstrapping the whole two-stage procedure. We use the non-parametric bootstrap with 200 replications. Bootstrapping a series of two models, the first of which is itself estimated by maximum-simulated likelihood, runs into serious constraints of computing time. Due to these constraints, we estimate a simplified version of the GMNL model that ignores scale heterogeneity, the mixed logit model (equation (2)) with a smaller number of random coefficients. We specify as random only the variables where every level of the variable had a significant standard deviation in the original GMNL model (only the on-call, location and earnings/income coefficients are specified as random) with other coefficients ‘fixed’. For each sample (of 1698 doctors) drawn with replacement, we estimate the mixed logit model (equation (2)), generate individual level-coefficients for the earnings attribute (equation (4)), and run the fixed effects regressions of GPs pricing decisions (equation (6)).

6. Results

Table A1 in the appendix shows the coefficient estimates from the first stage of the analysis – the GMNL model with a lognormally distributed earnings coefficient. This GMNL model was the best fit of the data according to BIC in comparison to both an MNL and MIXL model and compared to a GMNL model where the coefficient on earnings was normally distributed. The statistically significant value of τ in the GMNL model suggests the presence of scale heterogeneity, in addition the taste heterogeneity captured by the random coefficients. Figure 2 shows the distribution of the individual GP-level standardised marginal utility of income generated from this model according to the procedure outlined in equations (4) and (5).

We present descriptive statistics on standardised estimates of MUY and the characteristics of each GP in Table 1. Table 2 shows the association between the estimated marginal utility of income and characteristics of each GP in a linear regression. This estimation sample is larger (n=2,732) than that used in the competition model (n=1,698) due to item non response in the pricing variables which produce missing values in the competition model. Our results show that GPs aged over 65 years old have a higher marginal utility of income (0.32 of a standard deviation), than those aged under 65. Models including alternative specifications of age (linear, age squared, and 10 year age bands) had lower F-statistics and suggested that it is only the oldest age groups that demonstrate an association with the marginal utility of income. Seven percent of the estimation sample (n=183) are aged over 65 and their delayed retirement may indicate the need to earn more income. We also used experience as an alternative to age, with a similar result for those with more than 40 years of experience, but this model had a lower F-statistic.

The results also show GPs with higher annual household incomes are more likely to have a lower marginal utility of income. This effect is nonlinear, and is statistically significant only for those with incomes above the median annual household income of $208,500. Those above the median household income have marginal utility of income that is between 0.16 and 0.28 standard deviations lower than those in the lowest 10% of the distribution of household income (less than $100,000). This result is stable when the working spouse variable is excluded, as one would expect household income and partner working to be correlated.

GPs who qualified from Australian medical schools have a lower marginal utility of income, by 0.22 of a standard deviation, compared to those who qualified in non-Australian medical schools. There are also associations with current family circumstances. Those with dependent children or a non-working spouse have a higher marginal utility of income.

Table 3 shows descriptive statistics for the estimation sample used in the analysis of competition. The average price charged to all patients is almost AUD$42 and to non-bulk-billed patients is AUD$50. Note in the estimations we use the natural log of the two alternative price variables. The third dependent variable is the proportion of patients bulk-billed (charged zero co-payment). An average of 61% of patients are bulk-billed. The primary measure of competition is the distance to the third closest GP practice which has a mean of 1.51km. Again, we use the natural log of distance in the estimations.

Table 4 shows the key results from 12 regression models, three alternative pricing variables over four model specifications based on equation (6). Coefficients on log distance for the two price models can be interpreted as elasticities. In the bulk-billing model the coefficient is the percentage point change in the bulk-billing rate in response to a unit change in each explanatory variable. The coefficient on the marginal utility of income represents a one standard deviation increase in the marginal utility of income. The coefficients of interest are the direct effect of the MUY on prices, and the interaction term between the marginal utility of income and the log of distance.

The coefficients on the direct effect of distance to third closest competing GP practice mirror the results in Gravelle et al (2016). While Gravelle et al find an elasticity of 0.017 for the average price model, we find 0.015. The difference is caused by the inclusion of the MUY variables and the smaller estimation sample in our model (1698 vs 1966), otherwise the models are identical.

The estimated coefficient for the direct effect of MUY is always small and never statistically significant across the alternative models. The interaction coefficient is statistically significant in six of the 12 models. For the average price to all patients, the interaction is significant at the 1% level only in the area fixed effects model. When examining the prices to non-bulk-billed patients, the interaction is only significant in the fixed effects model at the 5% level. In the bulk-billing model, the interaction term is significant in all models, with the strongest effect in the fixed effects model.

The direction of the effect is as expected. GPs facing higher distances to their competitors and who have a higher marginal utility of income are more likely to charge higher prices compared to those with a lower marginal utility of income. For example, in the area fixed effects model for the average price to all patients, GPs with a marginal utility of income one standard deviation higher than the mean, have a distance elasticity almost double (0.026) that of the average of 0.015. With an average price of $41.98, this is equivalent to a price change of $0.63 for the average GP, and $1.09 for GPs who have a one standard deviation higher marginal utility of income. In the bulk billing model, the impact of a higher distance on bulk billing rates for GPs with a marginal utility of income one standard deviation higher than the mean, is a -4.98 percentage point reduction in the bulk-billing rate, almost double that of the average GP (-2.97).

An example of the full regression results for the area fixed effects model is shown in Table 5, with and without the marginal utility of income and its interaction. Table 5 also presents a robustness check where the MU_Y is estimated using the DCE data with the 'status quo' option, as opposed to the forced choice. This model is estimated with a more parsimonious specification of random parameters to aid convergence.[3] The results of the robustness check show the interaction term coefficient estimate retains the same sign and order of magnitude but is slightly smaller and, with a larger standard error, is no longer statistically significant at conventional levels. This result is consistent with the expectation that the model using the 'status quo' choice data estimates the MU_Y with less precision as there is less variation in the data. The first stage choice

[3] The model with the 'status quo' option data is estimated with only six random coefficients as opposed to twelve random coefficients in the models estimated on the 'forced choice' data

model results for the robustness check are presented alongside those for the main specification in Table A1.

Table 6 presents estimates of the area fixed effects models from Tables 4 and 5, alongside equivalent model estimates with bootstrapped standard errors as described at the end of section 5. The bootstrapped models have slightly different coefficient estimates as well as different standard errors due to the simplified mixed logit model used to estimate the MUy variable in the first stage. The models with bootstrapped standard errors have slightly larger standard errors for the coefficients on MUy and the interaction between MUy and log distance. However, the results are still statistically significant and very similar to those without bootstrapped standard errors.

7. Discussion

Heterogeneity in responses to financial incentives aimed at health care providers depends in part on differences in the relative weight they place on profit. A high degree of altruism and pro-social motivation may mean that some health care providers are less likely to change their behaviour in response to economic incentives. Though a feature of theoretical models of physician behaviour for decades, what is missing is empirical evidence on how motivation modifies the impact of economic incentives. As the public and private health care sector continues to roll out pro-competitive policies and pay for performance schemes for physicians and hospitals, evidence on their effects is mixed. This paper finds that the degree of a physician's monetary motivation influences their response to economic incentives.

Specifically, our results show the effect of competition on the prices charged by GPs in Australia depends on GPs' marginal utility of income. GPs with higher marginal utility of income and who are in less competitive areas, are more likely to charge higher prices compared to GPs in the same areas with a low marginal utility of income. We find particularly strong effects on the proportion of patients GPs choose to bulk bill, suggesting that this is the key margin on which they price discriminate.

We use a unique dataset that contains a DCE which was used to measure the marginal utility of income for each responding GP. Data were also gathered on prices charged by GPs, and the degree of competition was measured by distances to neighbouring general practices. We account for the endogeneity of distance and of the marginal utility of income using area fixed effects. These fixed effects accounted for unobservable characteristics of the local area that influence GPs location decisions and prices charged.

There is evidence that the marginal utility of income varies depending on personal, family and financial circumstances. GPs who qualified in an Australian medical school, and GPs with a working spouse, have a lower marginal utility of income. Marginal utility of income is higher for GPs with dependent children, for GPs aged over 65 years old, and those with lower household incomes. One might therefore expect that these groups of GPs will react differently to financial incentives. Changes in family composition, immigration policy, and eligibility for retirement may alter the motivation of segments of the GP workforce in Australia and could change the responsiveness of these GPs to financial incentives.

Our study does not directly measure altruism or pro-social motivation or examine the direct trade-off with the marginal utility of income. Whether one can assume that those with a low monetary motivation have a correspondingly high level of pro-social motivation depends on what other objectives they may have, and the nature of the relationship between them. Our results might also reflect the degree of intrinsic motivation of GPs. Those with a low monetary motivation may also have low extrinsic motivation and high intrinsic motivation, which have been shown to influence responses to the public reporting of performance and pay for performance schemes in health care (Kolstad 2013).

A key assumption in our analysis is that the marginal utility of income from the DCE reflects the GP's 'underlying' marginal utility of income, specifically that the marginal utility of income from choosing between jobs in the DCE is highly correlated to the marginal utility of income when deciding how to react to a change in competition. The significant effects found in our estimated models appear to validate this approach.

There is some support for the existence of a pre-determined level of marginal utility of income from neurophysiology. This literature links economic theories of rewards with behaviours by observing brain activity (Fehr and Camerer 2007, Schultz 2006). Money rewards have consistently been shown to activate neurons (reward circuits) in the prefrontal cortex of the brain, along with other types of rewards (O'Doherty et al. 2001, Knutson et al. 2005). As well as the pursuit of goals being undertaken consciously, some studies suggest that the response to monetary rewards can be unconscious or subliminal: in a given context a subject is unaware of the motivation but the brain still responds and pursues the reward (Pessiglione et al. 2007, Custers and Aarts 2010, Bijleveld, Custers, and Aarts 2011). This in turn suggests that the extent of monetary motivation is pre-existing in the brain, and individuals may unconsciously pursue a goal and invest effort in obtaining it. This, *"relies on associations between the representations of outcomes and positive reward signals that are shaped by one's history (for example, when a person was happy when making money or performing well). In this case, the goal is said to pre-exist as a desired state in the mind."* (Custers and Aarts 2010). If the goal representation is 'primed' unconsciously by a subliminal message or a factor associated with the goal or its representation, then an individual can pursue the goal without being aware that the goal is being pursued. There is also emerging research that the conscious awareness of the pursuit of monetary goals may be counter-productive in some circumstances (Bijleveld, Custers, and Aarts 2011).

This literature therefore suggests that the extent of monetary motivation may be pre-determined at the time at which a decision is made. This does not rule out that life events cannot influence the level of monetary motivation – indeed we have presented some evidence of this. Further research would be useful to examine how life events, particularly related to income shocks and altruistic goals, influence the marginal utility of income. If the marginal utility of money is pre-determined, then that also has implications for the causal effect of monetary motivation on observed behaviours, where at least reverse causality can be ruled out. The neurophysiology literature also suggests that monetary and other goals can be pursued unconsciously, suggesting again that there is an inbuilt level of monetary motivation that influences decisions without conscious knowledge of such activity.

The data we use are cross-sectional and the measure of marginal utility of income is from the same survey that also measured prices and earnings. In the future, longitudinal data can be used to examine the relationship between the marginal utility of income and earnings growth, and changes in competition. It will also be possible to repeat the choice experiment and examine whether the marginal utility of income has changed over time, and what factors might influence such changes.

When policy makers introduce financial incentives, they do not differentiate between different types of physician when designing incentives. There is a one size fits all approach. However, this may result in such incentives being ineffective on average, or only effective for certain sub-groups of physicians. Examining heterogeneity in response to financial incentives across observable characteristics remains an important area for future research. Examining how responses to competition vary according to the marginal utility of income of each physician, can potentially provide more accurate predictions of how doctors might react to changes in incentives in the future.

References

Arrow, Kenneth J. 1963. "Uncertainty and the Welfare Economics of Medical Care." *The American Economic Review* 53 (5):941-973.

Baltagi, BH, E Bratberg, and TH Holmas. 2005. "A panel data study of physicians' labour supply: the case of Norway." *Health Economics* 14 (10):1035-1045.

Besley, Timothy, and Maitreesh Ghatak. 2005. "Competition and incentives with motivated agents." *American Economic Review* 95 (3):616-636.

Bijleveld, Erik, Ruud Custers, and Henk Aarts. 2011. "Once the money is in sight: Distinctive effects of conscious and unconscious rewards on task performance." *Journal of Experimental Social Psychology* 47 (4):865-869. doi: 10.1016/j.jesp.2011.03.002.

Brekke, KR, H Gravelle, L Siciliani, and OR Straume. 2014. "Patient Choice, Mobility and Competition Among Health Care Providers." In *Health Care Provision and Patient Mobility*, edited by R Levaggi and M Montefiori. Springer-Verlag Italia.

Brekke, Kurt R., Luigi Siciliani, and Odd Rune Straume. 2011. "Hospital Competition and Quality with Regulated Prices." *Scandinavian Journal of Economics* 113 (2):444-469. doi: 10.1111/j.1467-9442.2011.01647.x.

Chomitz, KM, G Setiadi, A Azwar, M Ismail, and Widiyarti. 1998. What do doctors want? Developing Incentives for Doctors to Serve in Indonesia's Rural and Remote Areas. In *World Bank Policy Research Working Paper #1888*. Washingtion DC: World Bank.

Clemens, Jeffrey, and Joshua D Gottlieb. 2014. "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?" *The American Economic Review* 104 (4):1320.

Custers, Ruud, and Henk Aarts. 2010. "The Unconscious Will: How the Pursuit of Goals Operates Outside of Conscious Awareness." *Science* 329 (5987):47-50. doi: 10.1126/science.1188595.

Department of Health and Ageing. 2007. *Medicare Benefits Schedule. 1 Novermber 2007, Effective July 2008*. Canberra: Australian Government.

Ekeland, I, JJ Heckman, and L Nesheim. 2002. " Identifying hedonic models." *American Economic Review* 92 (2):304-309.

Ellis, Randall P, and Thomas G McGuire. 1986. "Provider behavior under prospective reimbursement: Cost sharing and supply." *Journal of Health Economics* 5 (2):129-151.

Ellis, RP, and TG McGuire. 1990. "Optimal payment systems for health services." *Journal of Health Economics* 9:375-396.

Evans, RG. 1984. *Strained mercy: the economics of Canadian health care*. Toronto: Butterworths.

Fehr, E, and CF Camerer. 2007. "Social neuroeconomics: the neural circuitry of social preferences." *Trends in Cognitive Sciences* 11 (10):419-427. doi: 10.1016/j.tics.2007.09.002.

Fehr, E, and A Falk. 2002. "Psychological foundations for incentives." *European Economic Review* 46:687-724.

Feldstein, M. 1970. "The rising price of physicians' services." *Review of Economics and Statistics* May:121-133.

Fiebig, Denzil G, Michael P Keane, Jordan Louviere, and Nada Wasi. 2010. "The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity." *Marketing Science* 29 (3):393-421. doi: 10.1287/mksc.1090.0508.

Frey, B. 1997. *Not Just for the Money: An Economic Theory of Personal Motivation*: Edward Elgar Publishing.
Frey, BS, and R Jegen. 2001. "Motivation Crowding Theory." *Journal of Economic Surveys* 15 (5):589-611.
Gaynor, M, and R Town. 2012. "Competition in health care markets." In *Handbook of Health Economics*, edited by Pauly M, McGuire T and Barros PP. Amsterdam: Elsevier.
Godager, Geir, and Daniel Wiesen. 2013. "Profit or patients' health benefit? Exploring the heterogeneity in physician altruism." *Journal of Health Economics* 32 (6):1105-1116. doi: http://dx.doi.org/10.1016/j.jhealeco.2013.08.008.
Gosden, T, I Bowler, and M Sutton. 2000. "How do general practitioners choose their practice? Preferences for practice and job characteristics." *Journal of Health Services Research and Policy* 5 (4):208 - 213.
Gosden, T, F Forland, IS Kristiansen, M Sutton, B Leese, A Giuffrida, M Sergison, and L Pedersen. 2001. "Impact of payment method on behaviour of primary care physicians: a systematic review." *Journal of Health Services Research and Policy* 6 (1):44-55.
Gravelle, Hugh, Anthony Scott, Peter Sivey, and Jongsay Yong. 2016. "Competition, prices and quality in the market for physician consultations." *The Journal of Industrial Economics* 64 (1):135-169. doi: 10.1111/joie.12098.
Greene, WH. 2007. *NLOGIT Version 4.0. Referecne Guide.* New York.: Econometric Software Inc.
Gregg, Paul, Paul A. Grout, Anita Ratcliffe, Sarah Smith, and Frank Windmeijer. 2011. "How important is pro-social behaviour in the delivery of public services?" *Journal of Public Economics* 95 (7–8):758-766. doi: http://dx.doi.org/10.1016/j.jpubeco.2011.03.002.
Hanson, Kara, and William Jack. 2010. "Incentives Could Induce Ethiopian Doctors And Nurses To Work In Rural Settings." *Health Affairs* 29 (8):1452-1460. doi: 10.1377/hlthaff.2009.0164.
Hennig-Schmidt, Heike, Reinhard Selten, and Daniel Wiesen. 2011. "How payment systems affect physicians' provision behaviour--An experimental investigation." *Journal of Health Economics* 30 (4):637-646. doi: 10.1016/j.jhealeco.2011.05.001.
Hole, AR, and JR Kolstad. 2010. Mixed logit estimation of willingness to pay distributions: a comparison of models in preference and WTP space using data from a health-relatedn choice experiment. In *Working Paper 03/10*. Bergen: Department of Economics, University of Bergen.
Iversen, Tor, and Hilde Lurås. 2000. "Economic motives and professional norms: the case of general medical practice." *Journal of Economic Behavior & Organization* 43 (4):447-470. doi: 10.1016/s0167-2681(00)00130-x.
Joyce, CM, A Scott, S-H Jeon, J Humphreys, G Kalb, J Witt, and A Leahy. 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* 10:50. doi: doi:10.1186/1472-6963-10-50.
Knutson, Brian, Jonathan Taylor, Matthew Kaufman, Richard Peterson, and Gary Glover. 2005. "Distributed Neural Representation of Expected Value." *The Journal of Neuroscience* 25 (19):4806-4812. doi: 10.1523/jneurosci.0642-05.2005.
Kolstad, Jonathan T. 2013. "Information and Quality When Motivation Is Intrinsic: Evidence from Surgeon Report Cards." *American Economic Review* 103 (7):2875-2910.

Kolstad, Julie Riise. 2011. "How to make rural jobs more attractive to health workers. Findings from a discrete choice experiment in Tanzania." *Health Economics* 20 (2):196-211. doi: 10.1002/hec.1581.

Kolstad, Julie Riise, and Ida Lindkvist. 2013. "Pro-social preferences and self-selection into the public health sector: evidence from an economic experiment." *Health Policy and Planning* 28 (3):320-327. doi: 10.1093/heapol/czs063.

Kruk, ME, JC Johnson, M Gyakobo, P Peter Agyei-Baffour, K Kwesi Asabir, SR Kotha, J Kwansah, E Nakua, RC Snow, and M. Dzodzomenyo. 2010. "Rural practice preferences among medical students in Ghana: a discrete choice experiment." *Bulletin of the World Health Organisation* 88:333-341. doi: 10.2471/BLT.09.072892.

Labelle, Roberta, Greg Stoddart, and Thomas Rice. 1994. "A re-examination of the meaning and importance of supplier-induced demand." *Journal of Health Economics* 13 (3):347-368. doi: 10.1016/0167-6296(94)90036-1.

Lagarde, M, and D Blaauw. 2009. "A review of the application and contribution of discrete choice experiments to inform human resources policy interventions." *Human Resources for Health* 7 (1):62.

Lagarde, Mylene, and Duane Blaauw. 2014. "Pro-social preferences and self-selection into jobs: Evidence from South African nurses." *Journal of Economic Behavior & Organization* 107, Part A (0):136-152. doi: http://dx.doi.org/10.1016/j.jebo.2014.09.004.

Lancsar, Emily, and Jordan Louviere. 2008. "Conducting Discrete Choice Experiments to Inform Healthcare Decision Making." *PharmacoEconomics* 26 (8):661-677.

Louviere, Jordan J., and Emily Lancsar. 2009. "Choice experiments in health: the good, the bad, the ugly and toward a brighter future." *Health Economics, Policy and Law* 4 (04):527-546. doi: doi:10.1017/S1744133109990193.

Manski, CF. 1977. "The structure of random utility models." *Theory and Decision* 8:229-254.

McGuire, T. 2000. "Physician Agency." In *Handbook of Health Economics*, edited by Newhouse J and Culyer AJ. Amsterdam: Elsevier.

Mooney, G., and M. Ryan. 1993. "Agency in health care: getting beyond first principles." *Journal of Health Economics* 12 (2):125.

O'Doherty, J, ML Kringelbach, ET Rolls, J Hornak, and C Andrews. 2001. "Abstract reward and punishment representations in the human orbitofrontal cortex." *Nature Neuroscience* 4 (95 - 102). doi: 10.1038/82959.

Pessiglione, Mathias, Liane Schmidt, Bogdan Draganski, Raffael Kalisch, Hakwan Lau, Ray J. Dolan, and Chris D. Frith. 2007. "How the Brain Translates Money into Force: A Neuroimaging Study of Subliminal Motivation." *Science* 316 (5826):904-906. doi: 10.1126/science.1140459.

Rizzo, John A., and Richard J. Zeckhauser. 2003. "Reference Incomes, Loss Aversion, and Physician Behavior." *Review of Economics and Statistics* 85 (4):909-922. doi: http://www.mitpressjournals.org/loi/rest.

Rizzo, John A., and Richard J. Zeckhauser. 2007. "Pushing incomes to reference points: Why do male doctors earn more?" *Journal of Economic Behavior & Organization* 63 (3):514-536. doi: 10.1016/j.jebo.2005.06.007.

Schultz, Wolfram. 2006. "Behavioral Theories and the Neurophysiology of Reward." *Annual Review of Psychology* 57 (1):87-115. doi: doi:10.1146/annurev.psych.56.091103.070229.

Scott, A, P Sivey, D Ait Ouakrim, L Willenberg, L Naccarella, J Furler, and D Young. 2011. "The effect of financial incentives on the quality of health care provided by primary care physicians." *Cochrane Database of Systematic Reviews* Issue 4 (Art. No.: CD008451). doi: DOI: 10.1002/14651858.CD008451.

Scott, A., J. Witt, J. Humphreys, C. Joyce, G. Kalb, S. H. Jeon, and M. McGrail. 2013. "Getting doctors into the bush: General Practitioners' preferences for rural location." *Social Science & Medicine* 96:33-44. doi: 10.1016/j.socscimed.2013.07.002.

Scott, Anthony. 2001. "Eliciting GPs' preferences for pecuniary and non-pecuniary job characteristics." *Journal of Health Economics* 20:329-347.

Serra, Danila, Pieter Serneels, and Abigail Barr. 2011. "Intrinsic motivations and the non-profit health sector: Evidence from Ethiopia." *Personality and Individual Differences* 51 (3):309-314.

Siciliani, Luigi. 2009. "Paying for performance and motivation crowding out." *Economics Letters* 103 (2):68-71. doi: 10.1016/j.econlet.2009.01.022.

Smith, Richard, Mylene Lagarde, Duane Blaauw, Catherine Goodman, Mike English, Kethi Mullei, Nonglak Pagaiya, Viroj Tangcharoensathien, Ermin Erasmus, and Kara Hanson. 2012. "Appealing to altruism: an alternative strategy to address the health workforce crisis in developing countries?" *Journal of Public Health*. doi: 10.1093/pubmed/fds066.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation*. 2nd ed. New York: Cambridge University Press.

Ubach, C, A Scott, F French, M Awramenko, and G Needham. 2003. "What do hospital consultants value about their jobs? A discrete choice experiment." *British Medical Journal* 326:1432.

Vujicic, M, M Alfano, B Shengelia, and S. Witter. 2010. Attracting doctors and medical students to rural Vietnam: insights from a discrete choice experiment. In *HNP Discussion Papers*. Washington, District of Columbia: World Bank.

Wordsworth, Sarah, Diane Skåtun, Anthony Scott, and Fiona French. 2004. "Preferences for general practice jobs: a survey of principals and sessional GPs." *British Journal of General Practice* 54:740-746.

Figure 1. Choice context and attributes.

Please read the following:

- You are asked to state which of two jobs (A or B) is better.
- You are then asked which job you would choose, including the option of staying in your current job.
- Everything about the jobs you are comparing is the same, except for the characteristics shown below.

Please use the following table to answer questions 5 and 6:

	Job A	Job B
Change in earnings	15% increase	No change
Change in total hours worked	No change	10% increase
On-call arrangements	1 in 4, frequently called out	1 in 2, frequently called out
Location	Coastal town, population < 5,000	City or large regional centre, population > 20,000
Social interactions	Very limited	Very good
Arranging a locum on short notice is	Rather difficult	Moderately easy
The practice team includes	GPs, receptionist and nurse	GPs, receptionist, nurse and practice manager
Average consultation length	15 minutes	> 20 minutes

5. Which job do you think is better? Job A ☐ Job B ☐
6. Which job would you choose? Job A ☐ Job B ☐ Stay at my current job ☐

Figure 2. Distribution of standardised marginal utility of income (kernel density)

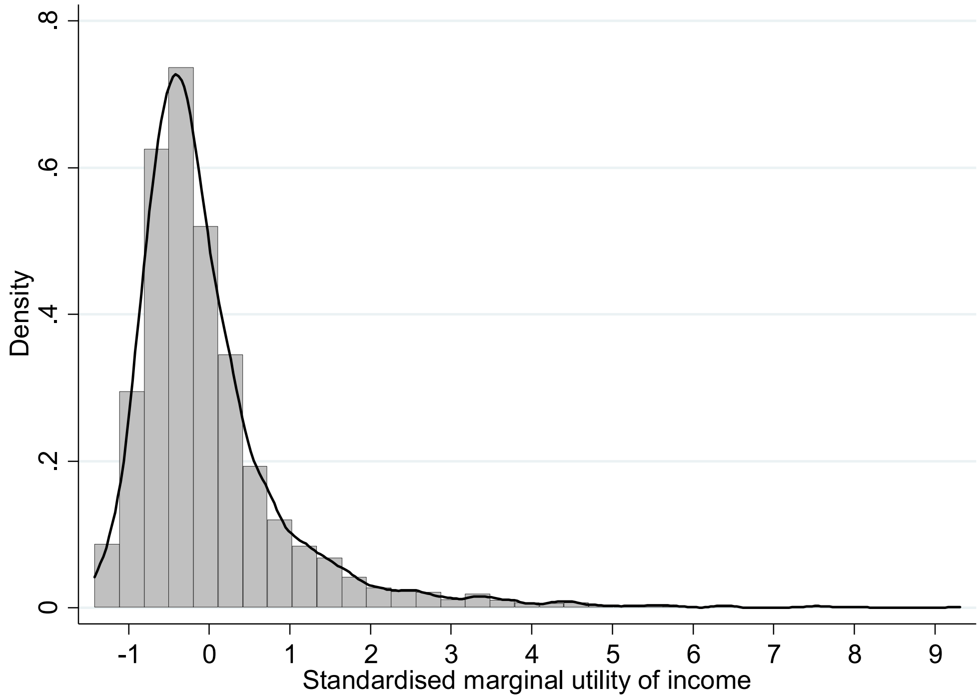

Table 1. Descriptive Statistics for estimation sample in Table 2 (n=2732)

	Mean	**sd**	**Min**	**Max**
Standardised MU_Y	0	1	-1.37	7.62
Female	0.45	0.5	0	1
Dependent children	0.67	0.47	0	1
Age	48.78	10.56	26	88
Aged over 65 years old	0.07	0.25	0	1
Australian medical School	0.78	0.41	0	1
Spouse not working	0.25		0	1
Spouse working	0.61		0	1
Not living with a spouse	0.14		0	1
	Mean	**Median**	**p25**	**p75**
Household income ($)	248,117	208,500	150,000	300,000

Table 2. Factors associated with the marginal utility of income[1]

	Coeff.	se
Aged over 65 years old (=1)	0.315	0.084
Household income (percentiles)[2]		
10% to 20%	-0.003	0.085
20% to 30%	-0.037	0.083
30% to 40%	-0.130	0.092
40% to 50%	-0.033	0.088
50% to 60%	-0.162	0.050
60% to 70%	-0.270	0.098
70% to 80%	-0.195	0.090
80% to 90%	-0.280	0.089
90% to 100%	-0.240	0.09
Australian Medical School (=1)	-0.222	0.046
Female (=1)	-0.020	0.041
Dependent children (=1)	0.102	0.045
Spouse working[3]	-0.106	0.048
Not living with a spouse/single[3]	0.122	0.067
Constant	0.122	0.084
Observations	2732	
F (15, 2716)	4.81	
R-squared	0.0259	

Notes: 1 = OLS regression: dependent variable is the standardised marginal utility of income from Figure 2. 2. 2 = omitted category is the bottom 10%. 3 = omitted category is spouse not working

Table 3. Descriptive Statistics for estimation sample in competition models (n=1698)

	Mean	SD	Min	Max
Dependent Variables:				
Av price all patients($): $m + (1-F^b)\ \bar{p}^{nb}$	41.916 49.973	9.077 11.057	41.916 49.973	9.077 11.057
Patients bulk-billed (%): F^b	60.854	31.266	60.854	31.266
Av price to non-bulk-billed($): $\bar{p}^{nb} + m$	41.916	9.077	41.916	9.077
Independent variables:				
Standardized MU_Y	0	1	-1.479	7.891
Third closest GP practice (km)	1.531	1.562	0.003	12.569
Ln(Third closest GP practice (km))	-0.002	0.985	-0.686	23.777
Female GP	0.476	0.500	0	1
Spouse	0.866	0.341	0	1
Children	0.655	0.475	0	1
Australian Medical School	0.814	0.389	0	1
Experience 10-19 years	0.219	0.414	0	1
Experience 20-29 years	0.376	0.484	0	1
Experience 30-39 years	0.256	0.436	0	1
Experience 40+ years	0.079	0.270	0	1
GP registrar	0.031	0.174	0	1
Partner or associate	0.453	0.498	0	1
Practice taxed as company	0.273	0.445	0	1
Practice size: 2-3 GPs	0.169	0.375	0	1
Practice size: 4-5 GPs	0.200	0.400	0	1
Practice size: 6-9 GPs	0.334	0.472	0	1
Practice size: 10+ GPs	0.161	0.367	0	1
SEIFA Index of adv/disadv	0	1	-4.521	2.242
Incentive area	0.230	0.421	0	1
Median House price ($0,000)	55.552	29.820	16.550	302.250
Proportion of residents U15	0.177	0.047	0.025	0.292
Proportion 65+	0.134	0.045	0.023	0.309
Proportion disabled	0.039	0.014	0.006	0.091
Proportion NW Europe	0.082	0.040	0.011	0.269
Proportion SE Europe	0.049	0.042	0.004502	0.300524
Proportion SE Asia	0.041	0.050	0.002424	0.421995
Proportion Other	0.095	0.080	0.002374	0.496274
Popn density (pop/km2) ('000)	2.020	1.587	0.019	8.757

Table 4. The effect of competition on prices

	OLS		Random effects		Mundlak		Area Fixed effects	
	Coeff	se	Coeff	se	Coeff	se	Coeff	se
Average price to all patients:								
Ln (3^{rd} closest practice km)	0.018	0.005	0.017	0.005	0.017	0.007	0.015	0.007
Standardized MU_Y	-0.005	0.004	-0.005	0.004	-0.005	0.004	-0.001	0.004
MU_Y*ln(3^{rd} closest practice km)	0.005	0.004	0.006	0.004	0.006	0.004	0.011	0.004
Price to non-bulk-billed patients:								
Ln (3^{rd} closest practice km)	0.020	0.007	0.020	0.007	0.019	0.011	0.019	0.011
Standardized MU_Y	0.001	0.006	0.001	0.006	0.000	0.006	0.005	0.006
MU_Y*ln(3^{rd} closest practice km)	0.002	0.007	0.003	0.007	0.003	0.006	0.013	0.006
Bulk-Billing rate:								
Ln (3^{rd} closest practice km)	-3.199	0.847	-3.016	0.802	-3.038	0.989	-2.97	1.019
Standardized MU_Y	0.648	0.687	0.632	0.704	0.705	0.710	0.126	0.778
MU_Y*ln(3^{rd} closest practice km)	-1.155	0.614	-1.19	0.615	-1.213	0.623	-2.008	0.638

Notes. For each dependent variable on the left, only the coefficients of distance and the marginal utility of income, and their interaction, are presented. All models includes full set of controls.

Table 5. The effect of competition on average price, detailed results (Area fixed-effects models)

	Without MU_Y		With MU_Y		With MUY (from status-quo DCE)	
	Coeff	**se**	**Coeff**	**se**	**Coeff**	**se**
Ln (3rd closest practice km)	0.016	0.007	0.015	0.007	0.019	0.007
Standardized MU_Y	-		-0.001	0.004	-0.001	0.006
MU_Y*ln(3rd closest practice km)	-		0.012	0.004	0.011	0.008
Female	0.039	0.011	0.038	0.011	0.039	0.011
Spouse	0.011	0.014	0.009	0.014	0.008	0.015
Dependent children	0.012	0.011	0.013	0.011	0.004	0.012
Australian medical school	0.061	0.013	0.060	0.012	0.061	0.013
Experience 10 to 19 yrs	0.045	0.021	0.046	0.021	0.052	0.022
Experience 20 to 29 yrs	0.033	0.021	0.035	0.022	0.039	0.021
Experience 30 to 39 yrs	0.047	0.023	0.049	0.023	0.054	0.023
Experience 40+ yrs	0.002	0.025	0.005	0.026	-0.008	0.026
Registrar	0.006	0.024	0.006	0.024	0.017	0.025
Partner	0.031	0.011	0.031	0.011	0.032	0.011
Company	-0.003	0.011	-0.003	0.011	0.001	0.011
Prac Size: 2-3 docs	-0.014	0.019	-0.013	0.019	-0.012	0.019
Prac Size: 4-5 docs	0.040	0.019	0.041	0.019	0.044	0.019
Prac Size: 6-9 docs	0.037	0.018	0.036	0.018	0.035	0.018
Prac Size: 10 or more	0.022	0.019	0.022	0.019	0.023	0.019
Constant	3.559	0.028	3.560	0.028	3.557	0.029
Observations	1698		1698		1627	
Number of groups	382		382		373	
F-statistic / Wald (df)	6.51 (16)		6.99 (18)		5.99 (18)	
R^2	0.077		0.078		0.0733	
Corr (u_i, $x\beta$)	0.064		0.058		0.043	

Table 6: Models with bootstrapped standard errors

	Area fixed effects		Area fixed effects bootstrapped	
	Coeff	**se**	**Coeff**	**se**
Average price to all patients:				
Ln (3rd closest practice km)	0.016	0.007	0.016	0.007
Standardized MU_Y	-0.001	0.004	-0.004	0.005
MU_Y*ln(3rd closest practice km)	0.013	0.004	0.011	0.005
Price to non-bulk-billed patients:				
Ln (3rd closest practice km)	0.019	0.011	0.019	0.009
Standardized MU_Y	0.008	0.005	0.001	0.005
MU_Y*ln(3rd closest practice km)	0.013	0.006	0.011	0.007
Bulk-Billing rate:				
Ln (3rd closest practice km)	-2.742	0.997	-3.089	1.128
Standardized MU_Y	0.062	0.677	0.186	0.853
MU_Y*ln(3rd closest practice km)	-2.413	0.634	-2.017	0.727

Notes. Each cell of three coefficient estimates and three standard errors represents a different model estimation. The first column of results reproduces the area fixed effects model estimates from Table 5. The second column of results presents equivalent estimates, but with standard errors bootstrapped with 200 replications. In this column, the first stage MUy estimates are produced from a simpler mixed logit model with a smaller number of random coefficients, therefore they also produce slightly different coefficient estimates in the second stage

Appendix 1.

Table A1. GMNL models used to recover individual-specific marginal utility of income

	Forced choice model		Status-quo model	
	Coefficient Mean	Coefficient SD	Coefficient Mean	Coefficient SD
Earnings ($‘000s)	-4.690 (0.112)	0.554 (0.172)	-4.543 (0.100)	0.782 (0.056)
Hours: 10% decrease	0.246 (0.016)		0.346 (0.032)	
Hours: 10% increase	-0.246 (0.020)		-0.330 (0.036)	
On call: 1 in 2	-1.137 (0.036)	0.605 (0.033)	-0.932 (0.045)	1.133 (0.034)
On call: 1 in 4, frequently	-0.082 (0.018)	0.089 (0.099)	-0.135 (0.042)	0.711 (0.036)
On call: 1 in 4, infrequently	0.585 (0.023)	0.128 (0.073)	0.146 (0.040)	0.896 (0.033)
Location: Inland, < 5,000	-0.225 (0.024)	0.506 (0.027)	-0.473 (0.044)	0.930 (0.032)
Location: Coastal, < 5,000	0.207 (0.022)	0.376 (0.030)	0.188 (0.043)	0.679 (0.032)
Location: Town, 5,000-20,000	-0.021 (0.022)		0.003 (0.031)	
Social interactions: Very limited	-0.318 (0.019)	0.364 (0.024)	-0.362 (0.025)	
Social interactions: Average	0.030 (0.018)		-0.057 (0.022)	
Arranging locum at short notice: Very difficult	-0.418 (0.019)	0.382 (0.026)	-0.303 (0.023)	
Arranging locum at short notice: Rather difficult	-0.032 (0.016)	0.161 (0.043)	-0.083 (0.026)	
Practice team: GP & receptionist	-0.272 (0.023)	0.399 (0.031)	-0.381 (0.033)	
Practice team: GP, rec. & nurse	0.007 (0.023)	0.146 (0.068)	-0.013 (0.045)	
Practice team: GP, rec., nurse & manager	0.108 (0.021)		0.177 (0.027)	
Consultation length: 10 min.	-0.421 (0.023)	0.405 (0.030)	-0.211 (0.030)	
Consultation length: 15 min.	0.084 (0.019)		0.103 (0.025)	
Consultation length: 20 min.	0.153 (0.021)		0.048 (0.028)	
Constant A	-0.029 (0.021)		-3.059 (0.055)	
Constant B	0.000 (0.000)		-3.200 (0.061)	
Tau	0.211 (0.000)		0.162 (0.000)	
Gamma	0.551 (0.033)		0.002 (0.013)	
Observations	31,905		28,719	
Individuals	1698		1627	
Log Likelihood	-17488		-8953	
BIC	35323		18204	
AIC	35043		17964	
Model Chi-sq	3635.9 (34 df)		40673.2 (29df)	

Notes: the coefficient of earnings is assumed to have a lognormal distribution, hence the estimated means and standard deviations of $\ln(\beta_{earn})$ are presented. Other coefficients with standard deviations are assumed to have normal distributions. For more detail on the methods of the DCE, see Scott et al (2013)