

Motivation and competition in health care

Anthony Scott¹  | Peter Sivey² 

¹Melbourne Institute: Applied Economic and Social Research, The University of Melbourne, Melbourne, Victoria, Australia

²Centre for Health Economics, University of York, York, UK

Correspondence

Anthony Scott, Melbourne Institute: Applied Economic and Social Research, The University of Melbourne, Melbourne, Victoria, Australia.

Email: a.scott@unimelb.edu.au

Funding information

National Health and Medical Research Council, Grant/Award Number: 454799; Australian Research Council, Grant/Award Number: FT0991296

Open access publishing facilitated by The University of Melbourne, as part of the Wiley - The University of Melbourne agreement via the Council of Australian University Librarians.

Abstract

Non-pecuniary sources of motivation are a strong feature of the health care sector and the impact of competitive incentives on behavior may be lower where pecuniary motivation is low. This paper measures the marginal utility of income (MUY) of physicians from a stated-choice experiment, and examines whether this measure influences the association between competition faced by physicians and the prices they charge. We find that physicians are more likely to exploit a lack of competition with higher prices if they have a high MUY.

KEYWORDS

competition, financial incentives, motivation, physicians

1 | INTRODUCTION

The effect of competition and other economic incentives in health care will depend, in part, upon the objectives of health care providers. The aim of this paper is to examine whether the degree of monetary motivation modifies the impact of economic incentives on physicians. This has not yet been examined empirically in health care. This paper examines whether physicians with a low monetary motivation, measured by their marginal utility of income (MUY), are less likely to respond to a change in competition relative to physicians with a relatively high monetary motivation. We contribute to the literature by linking preference parameters from a discrete choice experiment (DCE) to real world data on market structure and prices charged.

In health care, evidence that physicians respond to financial incentives is interpreted as physicians being motivated by income. However, theoretical models have always acknowledged that physicians may also be altruistic (Ellis & McGuire, 1986; Evans, 1984; Feldstein, 1970; Jack, 2005; Liu & Ma, 2013; McGuire, 2000; Siciliani, 2009), and combined with evidence that the effects of financial incentives can be mixed (Jia et al., 2021; Zaresani & Scott, 2021), suggests that there could be much heterogeneity in response to financial incentives that may be due, in part, to heterogeneity in the monetary motivation of physicians.

In general, if a physician places a relatively high weight on improving patients' health, it has been recognized that 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 (Besley & Ghatak, 2005; Mooney & Ryan, 1993). This contrasts with trends in health policy where physician remuneration schedules are becoming more complex over recent years with the introduction of pay

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. Health Economics published by John Wiley & Sons Ltd.

for performance schemes, suggesting that third-party payers think that physicians are motivated primarily by money. A key policy issue is how public sector organizations should design incentive schemes to influence behavior (Besley & Ghatak, 2003; Makris, 2009; Markis & Siciliani, 2013; Olivella & Siciliani, 2017). If people differ according to the source of their motivation, then policy makers need to consider these different 'types' of decision maker when designing policies that focus on the use of financial incentives versus, for example, social or group rewards.

Empirical work on the direct measurement of the motivation of health care providers, and its role in modifying the impact of financial incentives, is still in its infancy. Direct measurement of health professional's motivation has occurred in several contexts. This includes lab experiments of medical students (Attema et al., 2021; Godager & Wiesen, 2013; Hennig-Schmidt et al., 2011; Hennig-Schmidt & Wiesen, 2014; Kolstad & Lindkvist, 2013; Li, 2018; Smith et al., 2012; Wang et al., 2020) to measure altruism and pro-social preferences, with some studies for nurses linking these measures to actual behaviors outside of the lab (Kolstad & Lindkvist, 2013; Lagarde & Blaauw, 2014; Serra et al., 2011).

Other studies have focused on indirect measures of the monetary motivation of physicians. For example, Rizzo and Zeckhauser (2003) and Rizzo and Zeckhauser (2007) 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) led to reduction in the marginal utility of leisure and were found to provide more services to their patients.

The literature on DCEs is also a source of evidence on the contents of physician's utility functions and heterogeneity in their preferences for income. Respondents make choices between sets of alternative types of job, and job characteristics are varied according to an experimental design. An indirect utility function is estimated with parameters showing the marginal utility of each argument in the utility function, such as income and other job characteristics (Lagarde & Blaauw, 2009; Lancsar & Louviere, 2008; Louviere & Lancsar, 2009). Important to our context, DCEs provide direct estimates of the 'average' MUY across physicians (Gosden et al., 2000; Hanson & Jack, 2010; Kolstad, 2011; Kruk et al., 2010; Scott, 2001; Scott et al., 2013; Ubach et al., 2003; Vujcic et al., 2010; Wordsworth et al., 2004). Heterogeneity in MUY can be examined through interactions with observable physician characteristics such as income, age, gender, and experience (Chomitz et al., 1998; Hanson & Jack, 2010; Hole & Kolstad, 2010; Kruk et al., 2010; Scott, 2001; Ubach et al., 2003; Wordsworth et al., 2004).

This paper adds to the literature on competition in healthcare. Theoretical models consistently provide ambiguous predictions of the effects of competition on the behavior of healthcare providers, with effects depending on the details of demand and cost functions, as well as the payment scheme (Brekke et al., 2014; Gaynor & Town, 2012). These models also highlight that the effect of competition depends in part on the extent to which healthcare providers are motivated by altruism and income. Empirical research on competition in healthcare has focused largely on hospitals (Gaynor & Town, 2012), though there is a growing literature on competition between physicians (Bennett et al., 2015; Dietrichson et al., 2020; Dunn & Shapiro, 2014; Godager et al., 2015; Gravelle et al., 2016, 2019; Johar et al., 2014; Santos et al., 2015; Schaumans, 2015).

However, none of these studies have examined empirically the direct relationship between physicians' motivation and behavioral responses to competition.

The exceptions are two incentivized laboratory experiments, using the same experiment but on different samples (Byambadalai et al., 2021; Ge & Godager, 2021). They examine choices amongst different levels of quality involving direct trade-offs between profit and patient benefit under monopoly, duopoly, and quadropoly. However, though they used the framing of medical treatment choices, both used university student subjects from a range of disciplinary backgrounds, for example, with less than 10% studying medicine in Byambadalai et al. (2021), and so their findings are not generalizable to physicians. Both studies show that more competition increases quality, but their results are difficult to reconcile because each study uses a different structural model and different analytic methods, and reach different conclusions about why competition increases quality. Byambadalai et al. (2021) allow preferences to vary and conclude that more competition reduces altruism, whilst Ge and Godager (2021) keep preferences fixed and show that more competition reduces randomness in behavior and that this individual-specific unobserved heterogeneity (scale) is the main mechanism through which competition influences quality.

We add to this literature by examining the interaction between monetary motivation and competition using experimental measures of preferences applied to real world data on competition and prices. A simple Salop circular-city model of competition can show that greater distance between general practitioners (GPs) may lead to higher prices (Gravelle et al., 2016). In such a market where GPs have some market power and vary in their level of monetary motivation, we test two hypotheses. First, we test the hypothesis that GPs with a higher MUY will charge higher prices. Second, we test the hypothesis that there will be an interaction effect: GPs will respond to a fall in local competition by increasing prices more if they have a high MUY.

To account for potential vertical differentiation we also test these hypotheses with respect to a measure of GP service quality as well as price.

This hypothesis is examined in the Australian health care system which is a useful context to study competition between physicians. For 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 that has been used in two studies which we draw upon. The first study examined 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. Identification came from within area variation in prices and distances. We use the data and adapt the empirical model from Gravelle et al. (2016) to test whether the effect of competition is influenced by GPs' monetary motivation.

Our measure of monetary motivation is from a generalized multinomial logit model (G-MNL) (a form of mixed logit (MIXL) model) using data from a DCE administered to GPs in 2008 (Scott et al., 2013), from the same cross-sectional survey as in Gravelle et al. (2016). The DCE was used to estimate the marginal utility of an extra dollar of earnings for each GP through the estimation of a GMNL, after which it is possible to estimate the MUY for each individual GP. This estimated MUY GP is then interacted with the measure of competition in the empirical model estimated in Gravelle et al. The results suggest 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 MUY and present the results showing how the MUY 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. General practitioners are organized 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. General practitioners 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). There are four types of subsidy from Medicare available to patients 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 consultations¹ with a Medicare subsidy of A\$37.60 in 2018. The subsidy is slightly higher for children under 16 and for those with concession cards,² 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 enrollment.

General practitioners can charge what the market will bear – there are no fee controls or upper limits on fees charged. This means that there is a varying co-payment across different patients - the difference between the fee charged and the Medicare subsidy. This out of pocket payment cannot be covered by private health insurance according to legislation. Price discrimination can therefore occur in two ways. First, GPs can choose whether to 'bulk bill' for each service (i.e., visit) or charge a fee in excess of the subsidy. That is, to charge the patient only the Medicare subsidy so they face no out of pocket payment. Around 80% of GP visits are bulk billed.³ Some practices choose to bulk bill all patients, thereby competing with other practices directly on the basis of price. This does not necessarily mean that GPs in these practices have a lower monetary motivation, as these practices are usually for-profit companies and often provide short consultation lengths (i.e., low quality) and therefore a high volume of services per hour which can achieve high revenue. In other practices GPs have more discretion to price discriminate by choosing to bulk bill some patients but not others. This behavior may even reflect a lower monetary motivation, as GPs may bulk bill patients who are less affluent or with chronic conditions or with complex health and social problems. The direct effect of monetary motivation on bulk billing rates is therefore ambiguous, reflecting a high monetary motivation for those GPs in practices who bulk bill all patients (which also may have higher volume and lower quality of care), or a low monetary motivation for GPs who choose to bulk bill some patients but not others on the basis of patient's circumstances.

The second form of price discrimination is for patients who are not bulk billed, where different fees can be charged to different patients. However, within a group GP practice (only around 5% of practices are solo), there is less discretion on the size of the fee charged since the practice usually posts a fixed fee to be charged by all GPs in the practice, likely for administrative convenience, that is charged to all patients who visit that practice. The decision to change this fee therefore has higher adminis-

trative costs compared to deciding to bulk bill a patient or not for a specific visit which is up to the individual GP not the practice. Decisions to change the overall level of these fees are likely to be made less frequently. Evidence of price discrimination was found in Johar (2012) and Gravelle et al. (2016) where GPs located in affluent areas charged higher fees and were less likely to bulk bill. McIsaac et al. (2015) found that GPs are more likely to locate in affluent areas and that the health status of patients in areas (proxied by the mortality rate) does not influence location choices. These studies suggest that patient's ability to pay plays a role in both location and pricing decisions. Given that GPs are likely to have more discretion to bulk bill or not rather than to set fees, we expect that competition is more likely to influence bulk billing than fees for non-bulk billed patients. This result was found in Gravelle et al. (2016).

3 | 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% (3873/22,137) 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 (SLA) 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 382 SLAs with an average population of 33,998. 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.

The study was approved by the University of Melbourne Faculty of Business and Economics Human Ethics Advisory Group (Ref. 0,709,559) and the Monash University Standing Committee on Ethics in Research Involving Humans (Ref. CF07/1102–2,007,000,291).

4 | METHODS

We adopt a two-stage approach. First, we estimate a model of GPs preferences for job characteristics, including income, using data from the DCE in Scott et al. (2013). The empirical model is used to produce estimates of monetary motivation (the MUY) for each GP and examine its association with observable GP characteristics. Second, we incorporate our estimates of GPs' monetary motivation into the empirical model of competition and pricing decisions from Gravelle et al. (2016).

Scott et al. (2013) describe in detail the methods and results of the DCE which was included in Wave one of the MABEL survey in 2008 and completed by 3685 GPs. The DCE includes eight attributes. 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. This is so the MUY represents the marginal utility of a \$1 change in income rather than a one percent change in income. The other attributes are shown in Figure 1 and include hours worked (three levels), on-call arrangements (four levels), location (four

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 1 Example of the choice context and attributes [Colour figure can be viewed at wileyonlinelibrary.com]

levels), opportunities for social interactions (three levels), arranging a locum at short notice (three levels), practice team (four levels), and average consultation length (four levels).

Four attributes with three levels, and four with four levels gives a full factorial design of $3^4 \times 4^4 = 20,736$ possible scenarios. This is too many combinations to present to respondents, and so a fractional factorial experimental design was used to reduce the number of scenarios (Louviere et al., 2011). The experimental design varies the attribute levels to minimize the standard errors, and organizes the scenarios into choice pairs. Figure 1 shows an example of one of the eventual nine choice pairs each GP was asked to answer. The final design included 36 such choice pairs, each with two alternatives, blocked into four sets of nine choice pairs. General practitioners were randomly assigned to each block. Further details on the development of the DCE is provided in Scott et al. (2013). The experimental design was conducted in SAS, the DCE choice modeling used NLOGIT 5, and the final regression analysis was conducted in STATA 14.

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 maximize the amount of information about the preferences of each GP to obtain accurate estimates of the MUY. 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.

We estimate a G-MNL model, an extension of the MIXL model (Fiebig et al., 2010) which includes unobserved heterogeneity in marginal utilities through random parameters, providing an individual-specific estimate of the MUY. The distribution of the random parameters is given by the means and standard deviations of the random coefficients β_n (Train, 2009). The G-MNL model allows for scale heterogeneity, where the variance of the error terms varies across individuals (Hess & Train, 2017).⁴

For the purpose of this study, the important outputs from the G-MNL are the individual-level coefficient estimates (the MUY) which are available post-estimation. These are the expected values of the β_n given the parameter estimates and the choices made by each individual (Greene, 2007). Two GPs who completed the same set of nine choices, and choose the same alternatives, will therefore have the same individual-specific coefficient estimate of β_n (Train, 2009).

A log-normal distribution for the income coefficient is used 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 standardized 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 MUY estimated in the first stage and the response of GPs' pricing decisions to competition following the method of Gravelle et al. (2016). This paper examined the effect of competition on prices charged by GPs using data on self-reported fees and bulk billing from Wave 1 (2008) of the MABEL survey, and competition measured by the distance between GPs. The survey question about bulk billing is: "Approximately what percentage of patients do you bulk bill/charge no co-payment?" Fees are reported in the question "What is your current fee for a standard (Level B) consultation? (include Medicare rebate and patient co-payment if applicable. Please write dollar amount; write 0 if you bulk bill 100% of your patients." The total fee charged includes the Medicare rebate (m) and the out of pocket payment (p). The answers were used to construct three measures of price: (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 $\bar{p}^{nb} + m$ 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 $m + (1 - F^b) \bar{p}^{nb}$. The latter is calculated from a combination of the answers to the above two survey questions, and is an estimate of the average total price charged across patients who are and who are not bulk billed.

We also estimate a model with a proxy of quality as a dependent variable. The Gravelle et al. (2016) model predicts that more competition will increase quality. This is because more competition does not influence the quality provided to patients who are not bulk billed. But for patients who are bulk billed, an increase in competition leads to increases in quality since this is the only way to attract patients (as prices to these patients are already zero). General practitioners with a high monetary motivation are therefore more likely to increase quality compared to GPs with a low monetary motivation, but only for bulk billed patients. Following Gravelle et al. (2016) we use the reported average consultation time: "How long does an average consultation last (Please write number of minutes)."

We estimate linear models where measures of GP price are the dependent variable, and there are three key explanatory variables measuring competition (through distance to nearby competitors), the MUY (which we rename MUY) and the interaction of the two. This allows us to test the hypotheses that competition (GPdist) influences GP prices (β_1), that monetary motivation (through MUY) influences prices (β_{12}), and that the effect of competition is greater for GPs with a high monetary motivation (β_{11}):

$$y_{nr} = \beta_0 + \beta_1 \text{GPdist}_{nr} + \beta_{11} \text{GPdist}_{nr} * \text{MUY}_{nr} + \beta_{12} \text{MUY}_{nr} + \beta_2 \text{GPchars}_{nr} + \beta_3 \text{Areachars}_r + \epsilon_{nr} \quad (1)$$

The dependent variable y_{nr} is one of the three alternative measures of prices or the consultation time measure. 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 (SLA), 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 the population born overseas, and population density.

To measure competition, we use the measure from Gravelle et al. (2016) which is the straight-line road distances between the address of each GP survey respondent and nearby GP addresses from the population of GPs. This provides a GP-level measure of competition. For example, if there are two GPs within a small area they will be the same distance from each other, but each will have a different measure of competition as the nearest GP may be located in an adjacent small area. In the analysis presented here we use the third nearest GP practice, and conduct robustness checks with the distance to the nearest and fifth nearest GP practice. Use of the third nearest is based on results from Bresnahan and Reiss (1991) who show that only the first three additional competitors in a market have a large effect on prices in geographically isolated markets for health professional services. This individual measure of competition improves on area-based measures such as market shares of densities, since it allows the use of area fixed effects to account for unobservable factors influencing both prices and competition (Gravelle et al., 2016). A key issue in the previous literature is that competition is usually measured at a small area level and that small area characteristics (e.g., schools, availability of hospitals, amenities, practice costs) also influence the number of GPs in that area and therefore competition, leading to biased estimates.

Unobserved area-specific factors include the characteristics that influence both pricing decisions and physician's decisions to work in that area and therefore their distance from other GPs. This includes characteristics of the population (demand, preferences, health status, socio-economic status). These may reflect preferences for treating certain groups of patients, but also may reflect factors that influence expected profits, such as factors influencing costs of setting up a private practice, such

as property prices/rental costs, and access to amenities, or other nearby health services, such as hospitals, that can influence referral networks.

For area-specific unobservables, we exploit the fact that we have an individual-specific measure of competition (distance to the nearest other GP), and so can use area fixed effects to control for all unobserved factors that are the same within areas, and which are likely be correlated with the physician's choice to locate to that area and the physician's pricing decisions. Identification therefore comes from within area variation in prices, distances and MUY. So though the data are from a cross-section, they have a 'panel' structure (physicians within areas).

The second source of unobservables are those which are correlated with the MUY and prices. These include both area and physician characteristics. Unobserved area characteristics might attract GPs with a high or low MUY to certain areas. For example, affluent areas may attract GPs with a high MUY, as previous research has shown that GPs price discriminate and charge richer patients more (Gravelle et al., 2016; Johar, 2012), as might areas with low property prices/rental costs of setting up a practice. Our area fixed effects will capture these unobservables. Identification comes from within-area variation in the MUY, distances between practices, and prices charged by GPs.

There may also be GP characteristics that are correlated with both prices and the MUY. To the extent that the MUY is associated with life cycle and family factors of the GP, a rich set of observable GP characteristics are included in the set control variables in the regression models. To the extent that GPs with the same unobserved characteristics choose to work in the same areas, then our area fixed effects will capture these unobserved GP characteristics. However, there may remain some unobserved GP characteristics that vary within areas, and which are correlated with GPs' MUY and prices charged. The areas, with an average population of 33,998, are small enough that unobserved factors varying within areas and which are correlated with within area variation in prices and distance are likely to be negligible.

Overall, our identification strategy requires that GP location decisions within areas are uncorrelated with within-area varying factors affecting pricing decisions. This strategy is appropriate if doctors choose to practice in an area, such as a suburb of a city, while their exact location within the area is determined by factors which do not influence their pricing decisions, such as a vacancy at an existing GP practice, planning restrictions, or the availability of vacant premises for the setting up of a practice. We believe that these factors are plausibly exogenous and independent of GP pricing within local areas.

5 | RESULTS

We first document the factors associated with the MUY. This helps to assess the external validity of our measure from the DCE. The standardized estimate of the MUY for each GP (β_n) from the DCE is used as the dependent variable in a linear regression. Table A1 in the Appendix shows the coefficient estimates from the first stage of the competition analysis – the G-MNL model with a lognormally distributed earnings coefficient. This G-MNL model was the best fit of the data according to Bayesian Information Criteria (BIC) in comparison to both an MNL and MIXL model and compared to a G-MNL model where the coefficient on earnings was normally distributed. The statistically significant value of τ in the G-MNL model suggests the presence of scale heterogeneity, in addition to the taste heterogeneity captured by the random coefficients. This model was used to generate the estimates of MUY used in the competition analysis.

Descriptive statistics are in Table 1. Table 2 shows that GPs aged over 65 years old have a higher MUY (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 MUY. Seven percent of the estimation sample ($n = 183$) are aged over 65 and their delayed retirement may indicate the need to earn more income.

The results show GPs with higher annual household incomes are more likely to have a lower MUY. 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 MUY 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.

General practitioners who qualified from Australian medical schools have a lower MUY, 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 MUY.

A smaller sample of GPs is used ($n = 1698$) when the MUY estimates are matched to the sample used in the competition model from Gravelle et al. (2016). This is because the competition model only used data from GPs in metropolitan areas and requires respondents to have complete pricing information.

TABLE 1 Descriptive statistics for estimation sample in Table 2 ($n = 2732$)

	Mean	Sd	Min	Max
Standardized MU _y	0.00	1.00	-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

	Coeff.	se
Aged over 65 years old (=1)	0.315***	0.084
Household income (percentiles) ^b		
10%–20%	-0.003	0.085
20%–30%	-0.037	0.083
30%–40%	-0.130	0.092
40%–50%	-0.033	0.088
50%–60%	-0.162***	0.050
60%–70%	-0.270***	0.098
70%–80%	-0.195**	0.090
80%–90%	-0.280***	0.089
90%–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 ^c	-0.106**	0.048
Not living with a spouse/single ^c	0.122*	0.067
Constant	0.122*	0.084
Observations	2732	
F (15, 2716)	4.81***	
R-squared	0.0259	

^aOLS regression: dependent variable is the standardized marginal utility of income.

^bOmitted category is the bottom 10%.

^cOmitted category is spouse not working.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

TABLE 2 Factors associated with the marginal utility of income (MUY)^a

Table 3 shows descriptive statistics for the estimation sample used in the competition model. The average price charged to all patients is almost AUD\$42 and to non-bulk billed patients is close to AUD\$50. 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), which has a mean of 61%. The primary measure of competition is the distance to the third closest GP practice which has a mean of 1.51 km. Again, we use the natural log of distance in the estimations.

As our models rely on within-area variation in competition (distance between GP practices), prices (or consultation time) and MUY, Table 3 includes both within and between-area standard deviations in all GP-level variables. These figures demonstrate substantial within-area variation in all key variables with within-area variation in prices and MUY greater than or equal

TABLE 3 Descriptive statistics for the estimation sample in competition and prices models ($n = 1698$)

	Mean	SD	SD between-areas	SD within- areas	Min	Max
Dependent variables:						
Av price to non-bulk billed (\$): $\bar{p}^{nb} + m$	49.973	11.057	8.239	8.291	32.8	120
Patients bulk billed (%): F^b	60.854	31.266	24.667	23.491	0	100
Av price all patients (\$): $m + (1-F^b)\bar{p}^{nb}$	41.916	9.077	7.056	6.707	32.8	98.2
Independent variables:						
Standardized MUY	0.023	0.945	0.567	0.843	-1.479	7.891
Closest GP practice (km)	0.697	0.970	1.025	0.615	0.000	9.434
Third closest GP practice (km)	1.531	1.562	1.592	0.892	0.003	12.569
Fifth closest GP practice (km)	2.180	1.934	1.989	0.949	0.067	15.912
Female GP	0.476	0.500	0.342	0.432	0	1
Spouse	0.866	0.341	0.213	0.303	0	1
Children	0.655	0.475	0.309	0.418	0	1
Australian medical school	0.814	0.389	0.299	0.328	0	1
Experience 10–19 years	0.219	0.414	0.287	0.360	0	1
Experience 20–29 years	0.376	0.484	0.306	0.426	0	1
Experience 30–39 years	0.256	0.436	0.319	0.377	0	1
Experience 40+ years	0.079	0.270	0.149	0.244	0	1
GP registrar	0.031	0.174	0.093	0.166	0	1
Partner or associate	0.453	0.498	0.346	0.426	0	1
Practice taxed as company	0.273	0.445	0.311	0.378	0	1
Practice size: 2–3 GPs	0.169	0.375	0.299	0.318	0	1
Practice size: 4–5 GPs	0.200	0.400	0.300	0.330	0	1
Practice size: 6–9 GPs	0.334	0.472	0.344	0.385	0	1
Practice size: 10+ GPs	0.161	0.367	0.214	0.312	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	-	-	0.004	0.300
Proportion SE Asia	0.041	0.050	-	-	0.002	0.422
Proportion other	0.095	0.080	-	-	0.002	0.496
Popn density (pop/km2) ('000)	2.020	1.587	-	-	0.019	8.757

to between-area variation. Understandably, for the distance measures of competition, between-area variation is the larger, but within-area variation is still at least half as large as between-area variation.

Table 4 shows the key results from 16 regression models, three alternative pricing variables and consultation time, over four model specifications based on Equation (1). The preferred model is the model with area fixed effects in the last two columns. Coefficients on log distance for the two price models and the consultation time model 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. As the MUY is standardized to have a mean of zero and a standard deviation of 1, the coefficient on the MUY represents a one standard deviation increase in the MUY. The coefficients of interest are the direct effect of the MUY on prices, and the interaction term between the MUY and the log of distance.

TABLE 4 The effect of competition and monetary motivation on prices

	OLS		Random effects		Mundlak		Area fixed effects	
	Coeff	se	Coeff	se	Coeff	se	Coeff	se
Log average price to all patients:								
Ln (3 rd closest practice km)	0.018***	0.005	0.017***	0.005	0.016**	0.007	0.015**	0.007
Standardized MUY	-0.005	0.004	-0.005	0.004	-0.003	0.005	-0.001	0.004
MUY*Ln (3 rd closest practice km)	0.005	0.004	0.006	0.004	0.010**	0.004	0.012***	0.004
Log price to non-bulk billed patients:								
Ln (3 rd closest practice km)	0.021***	0.007	0.020**	0.007	0.019*	0.011	0.019*	0.011
Standardized MUY	0.001	0.006	0.001	0.006	0.004	0.006	0.005	0.006
MUY*Ln (3 rd closest practice km)	0.002	0.007	0.003	0.007	0.011	0.007	0.013**	0.006
Bulk billing rate:								
Ln (3 rd closest practice km)	-3.199***	0.847	-3.016***	0.802	-3.027***	1.001	-2.97**	1.019
Standardized MUY	0.648	0.687	0.632	0.704	0.343	0.773	0.126	0.778
MUY*Ln (3 rd closest practice km)	-1.155*	0.614	-1.194*	0.615	-1.830***	0.639	-2.008***	0.638
Log consultation time:								
Ln (3 rd closest practice km)	-0.004	0.009	-0.005	0.009	-0.013	0.012	-0.016	0.012
Standardized MUY	0.010	0.007	0.009	0.007	0.005	0.007	0.005	0.008
MUY*Ln (3 rd closest practice km)	-0.002	0.006	-0.002	0.006	-0.004	0.007	-0.003	0.007

Note: 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 (see Table 5). Sample size $n = 1698$ in all models.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

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 as not all GPs in Gravelle et al. (2016) completed the DCE (1698 vs. 1966).

For the pricing outcome variables, the area fixed effects models (last column) show the coefficients of the interaction terms are all statistically significant (at the 1% level for the average price to all patients and for the bulk billing rate, and at the 5% level for the price to non-bulk billed patients).

The direction of the effect is as expected. General practitioners facing longer distances to their competitors and who have a higher MUY are more likely to charge higher prices compared to those with a lower MUY. For example, in the area fixed effects model for the average price to all patients, GPs with a MUY one standard deviation higher than the mean (of zero), have a distance elasticity almost double (0.026) that of the baseline 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 MUY. In the bulk billing model, the impact of a higher distance on bulk billing rates for GPs with a MUY one standard deviation higher than the mean, is a -4.98% point reduction in the bulk billing rate, almost double that of the average GP (-2.97).

For the consultation time outcome variable, none of the coefficients measuring the effects of competition, the MUY, or their interaction, are statistically significant at conventional levels. This result suggests that neither GPs' MUY nor its interaction with competition have an effect on consultation time.

The estimated coefficient for the direct effect of MUY is always small and never statistically significant across the alternative models. The estimated effect of a one standard deviation change in the MUY on price, at the mean distance (1.53 km) and mean price (\$41.92) is \$0.15 and at a distance one standard deviation higher than the mean (3.09 km) is \$0.48. The reason for this small effect could be that other factors correlated with the MUY in Table 2, and that are included as independent variables in the main models in Table 4, are capturing the effect. This included being qualified in an Australian medical school and having dependent children.

The interaction terms are not statistically significant for the ordinary least squares (OLS) and random effects models, except for the bulk billing dependent variable. For the pricing dependent variables, across all specifications the size of interaction effect is larger in the Mundlak specification and largest in the area fixed effects models. This suggests the existence of unob-

TABLE 5 The effect of competition on log average price, detailed results (area fixed-effects models)

	Without MUY		With MUY		With MUY (from status-quo DCE)	
	Coeff	se	Coeff	se	Coeff	se
Ln (3 rd closest practice km)	0.016**	0.007	0.015**	0.007	0.019**	0.007
Standardized MUY	-		-0.001	0.004	-0.001	0.006
MUY*Ln (3 rd 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 10–19 years	0.045**	0.021	0.046**	0.021	0.052**	0.022
Experience 20–29 years	0.033	0.021	0.035*	0.022	0.039*	0.021
Experience 30–39 years	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 ($u_p, x\beta$)	0.064		0.058		0.043	

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

served area effects that lead to downward bias in the size of the interaction term for each of the price variables. An example of the full regression results for the area fixed effects model is shown in Table 5, with and without the MUY and its interaction.

6 | ROBUSTNESS CHECKS

Table 5 also presents a robustness check where the MUY is estimated using the DCE data with the ‘status quo’ option, as opposed to the forced choice. Recall (Figure 1) this included the option for GPs to state that they prefer to stay in their current job. We expect the model using the ‘status quo’ choice data estimates the MUY with less precision as there is less variation in the data since a higher proportion of people will choose their own job rather than job A or job B. This model is estimated with a more parsimonious specification of random parameters to aid convergence.⁵ 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 larger standard error is to be expected as the model using the “status quo” choice data estimates the MUY much less precisely. This lack of precision is because a high proportion of GPs choose the “status quo” option for many of their responses which provides less information for estimating preference parameters. The first stage choice model results for the robustness check are presented alongside those for the main specification in Table A1.

As the estimated MUY 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

TABLE 6 Models with bootstrapped standard errors

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

Note: Each cell of three coefficient estimates and three standard errors represents a different model estimation. Sample size $n = 1698$ in all models. 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 MU_Y 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.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

(Table 6). 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)). 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 the MUY and its interaction with log distance. However, the results are still statistically significant and very similar to those without bootstrapped standard errors.

Table 7 includes the same area fixed effects models from Tables 4 and 5 for the three different pricing outcome measures, for three alternative distance measures of competition. This tests whether different measures of competition change the results. While the models using the distance to the closest GP practice fail to attain statistical significance in the key results, the models using the distance to the fifth closest GP practice show the same pattern of results as when the distance to the third closest is used. These results follow the same pattern as in Gravelle et al. (2016) where the size of the effects are substantially higher using the fifth closest distance measure, possibly explained by the larger variation in this variable.

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 monetary motivation (and a low degree of other motivation, such as altruism) may mean that some health care providers are more likely to change their behavior in response to economic incentives. Though a feature of theoretical models of physician behavior for decades, what is missing is empirical evidence on how heterogeneity in 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.

TABLE 7 Models with alternative distance measures of competition (area fixed effects models)

	Distance measure: 3 rd closest practice (km)		Distance measure: Closest practice (km)		Distance measure: 5 th closest practice (km)	
	Coeff	Se	Coeff	se	Coeff	se
Log average price to all patients:						
Ln (Distance)	0.015**	0.007	0.002	0.004	0.032***	0.010
Standardized MU _Y	-0.001	0.004	0.001	0.006	-0.008	0.005
MUY*Ln (Distance)	0.011***	0.004	0.004	0.003	0.014***	0.005
Log price to non-bulk billed patients:						
Ln (Distance)	0.019*	0.011	0.004	0.005	0.040***	0.014
Standardized MU _Y	0.005	0.006	0.003	0.007	-0.002	0.008
MUY*Ln (Distance)	0.013**	0.006	0.000	0.004	0.017**	0.008
Bulk billing rate:						
Ln (Distance)	-2.97**	1.019	-0.521	0.595	-5.815***	1.553
Standardized MU _Y	0.126	0.778	-0.551	0.982	1.074	0.894
MUY*Ln (Distance)	-2.008***	0.638	-0.748	0.458	-2.029**	0.799

Note: Each cell of three coefficient estimates and three standard errors represents a different model estimation. Sample size $n = 1698$ in all models. The first column of results reproduces the area fixed effects model estimates from Table 5 with distance to the third closest practice measuring competition. The second column of results presents equivalent estimates, but with the distance measure replaced with the distance to the (first) closest practice in kilometers. The third column replaces the distance measure with the distance to the fifth closest practice in kilometers.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Specifically, our results show the effect of competition on the prices charged by GPs in Australia depends on GPs' MUY. General practitioners with higher MUY and who are in less competitive areas, are more likely to charge higher prices compared to GPs in the same areas with a low MUY. 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. While the coefficient on the direct effect of MUY is not statistically significant in our models, its direct effect on price is positive, although relatively small, when evaluating the interaction term at the mean of the distance measure of competition. The measure of MUY did not have a statistically significant direct effect, or interact with competition in affecting our proxy measure of consultation quality – consultation time. The competition measure does not have a statistically significant effect on quality as a direct effect either (in Gravelle et al., 2016), suggesting quality is not a margin for competition for Australian GPs.

Our study does not directly measure altruism or pro-social motivation or examine the direct trade-off with the MUY. 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).

Our measure of the MUY is taken from a stated preference experiment. Discrete choice experiments are not usually incentivized, raising questions about the external validity of our measure. External validity in DCEs is usually strengthened by the choice task being carefully designed and piloted, using qualitative research, to replicate a salient choice situation and ensure respondents understand what is being asked. A recent meta-analysis and systematic review of the external validity of DCEs found that they had moderate accuracy in predicting real behaviors (Quaife et al., 2018). Some evidence to support external validity is that in the DCE in 84.2% of choices, GPs preferred to stay in their current job with 15.8% stating a preference to move (Scott et al., 2013). Data from the population of GPs in Australia in the same year (2008/9) show that 10.8% moved jobs (Yong et al., 2018), providing accurate predictions around 68% of the time. This seems reasonable. In addition, the job characteristics were very familiar to the highly educated group of respondents. Even if they were not considering a move they would have reasonably well-formed preferences and information about their job characteristics therefore reducing the chance of bias. Furthermore, the choice task had few obvious reasons for strategic bias in GPs' responses.

External validity also rests on the assumption in our analysis that the MUY from the DCE reflects the GP's 'underlying' MUY. This is consistent with evidence from brain studies showing trade-offs between selfish and pro-social rewards (Fehr & Camerer, 2007). In particular, we assume that the MUY estimated from choosing between jobs in the DCE is highly correlated

to the MUY when deciding how to react to a change in competition. The significant effects found in our estimated models appear to validate this approach. Furthermore, in addition to the literature suggesting that personality traits are reasonably stable during adulthood (Cobb-Clark & Schurer, 2012), our analysis in Table 2 also shows that the MUY is associated with some GP characteristics, including life cycle factors. There is evidence that the MUY varies depending on personal, family and financial circumstances. General practitioners who qualified in an Australian medical school, and GPs with a working spouse, have a lower MUY. Marginal utility of income is higher for GPs with dependent children, for GPs aged over 65 years old, and those with lower household incomes. This provides some evidence of external validity. 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 could alter the motivation of segments of the GP workforce in Australia and could change the responsiveness of these GPs to financial incentives. Further research would be useful to examine how life events, particularly related to income shocks and altruistic goals, influence the MUY.

Our measure of competition is limited by the fact that we do not have information on the market shares of each GP, as our data are not linked to patients, and we cannot model patient choice of GPs as an identification strategy. Distances between providers are a plausible alternative that have been used before, and the use of individual-specific distances allow the use of area fixed effects to account for unobservables that are constant across a small area and correlated with both distances and prices.

When policymakers 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 approach 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 MUY of each physician, can potentially provide more accurate predictions of how doctors might react to changes in incentives in the future.

ACKNOWLEDGMENTS

Funding is acknowledged from an Australian Research Council Future Fellowship (FT0991296) and an National Health and Medical Research Council (NHMRC) Principal Research Fellowship (628,471). Used data from the MABEL longitudinal survey of doctors. Funding for MABEL was provided by the NHMRC (2007–2016: 454799 and 1019605); the Australian Department of Health and Aging (2008); Health Workforce Australia (2013); The University of Melbourne, Medibank Better Health Foundation, the New South Wales Department of Health, and the Victorian Department of Health and Human Services (2017); and the Australian Government Department of Health, the Australian Digital Health Agency, and the Victorian Department of Health and Human Services (2018). We thank the doctors who gave their valuable time to participate in MABEL, and the other members of the MABEL research team. Open access publishing facilitated by The University of Melbourne, as part of the Wiley - The University of Melbourne agreement via the Council of Australian University Librarians.

DATA AVAILABILITY STATEMENT

Medicine in Australia: Balancing Employment and Life survey data are available from the Australian Data Archive. These do not include the DCE data or the competition variables. The authors will consider requests for the data used in the estimation samples for the purposes of replication.

ORCID

Anthony Scott  <https://orcid.org/0000-0002-2851-5378>

Peter Sivey  <https://orcid.org/0000-0002-3703-615X>

ENDNOTES

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

² A concession card is held by the elderly and those on welfare payments that entitles them to certain discounts and subsidies for public services.

³ <http://www.health.gov.au/medicarestats> [accessed 07/10/2016]

⁴ Scale heterogeneity may exist due to near-lexicographic preferences where marginal utilities for some attributes are very high (i.e., scaled up); or at the other extreme can be due to randomness of behavior 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.

⁵ The model with the ‘status quo’ option data is estimated with only six random coefficients as opposed to 12 random coefficients in the models estimated on the ‘forced choice’ data

REFERENCES

- Attema, A. E., Galizzi, M. M., Gros, M., Hennig-Schmidt, H., Karay, Y., L'Haridon, O., & Wiesen, D. (2021). *The formation of physician altruism*. EsCHER working paper series No. 2021001. Erasmus University.
- Australian Institute of Health and Welfare. (2010). *Medical labour force 2008. Bulletin no. 82. Data iTble 3.2 Cat. no. AUS 131*. AIHW.
- Bennett, D., Hung, C.-L., & Lauderdale, T.-L. (2015). Health care competition and Antibiotic use in Taiwan. *The Journal of Industrial Economics*, 63(2), 371–393. <https://doi.org/10.1111/joie.12075>
- Besley, T., & Ghatak, M. (2003). Incentives, choice, and accountability in the provision of public services. *Oxford Review of Economic Policy*, 19(2), 235–249. <https://doi.org/10.1093/oxrep/19.2.235>
- Besley, T., & Ghatak, M. (2005). Competition and incentives with motivated agents. *The American Economic Review*, 95(3), 616–636. <https://doi.org/10.1257/0002828054201413>
- Brekke, K. R., Gravelle, H., Siciliani, L., & Straume, O. R. (2014). Patient choice, mobility and competition among health care providers. In R. Levaggi & M. Montefiori (Eds.), *Health care provision and patient mobility* (pp. 1–26). Springer-Verlag Italia. https://doi.org/10.1007/978-88-470-5480-6_1
- Bresnahan, T. F., & Reiss, P. C. (1991). Entry and competition in concentrated markets. *Journal of Political Economy*, 99(5), 977–1009. <https://doi.org/10.1086/261786>
- Byambadai, U., Ma, C., & Wiesen, D. (2021). *Changing preferences: An experiment and estimation of market-incentive effects on altruism*.
- Chomitz, K., Setiadi, G., Azwar, A., Ismail, M., & Widiyarti. (1998). *What do doctors want? Developing incentives for doctors to serve in Indonesia's rural and remote areas*. World Bank Policy Research Working Paper #1888. World Bank.
- Cobb-Clark, D. A., & Schurer, S. (2012). The stability of big-five personality traits. *Economics Letters*, 115(1), 11–15. <https://doi.org/10.1016/j.econlet.2011.11.015>
- Department of Health and Ageing. (2007). *Medicare benefits schedule. 1 November 2007, Effective July 2008*. Australian Government. Effective July 2008.
- Dietrichson, J., Ellegård, L. M., & Kjellsson, G. (2020). Patient choice, entry, and the quality of primary care: Evidence from Swedish reforms. *Health Economics*, 29(6), 716–730. <https://doi.org/10.1002/hec.4015>
- Dunn, A., & Shapiro, A. H. (2014). Do Physicians possess market power? *The Journal of Law and Economics*, 57(1), 159–193. <https://doi.org/10.1086/674407>
- Ellis, R. P., & McGuire, T. G. (1986). Provider behavior under prospective reimbursement: Cost sharing and supply. *Journal of Health Economics*, 5(2), 129–151. [https://doi.org/10.1016/0167-6296\(86\)90002-0](https://doi.org/10.1016/0167-6296(86)90002-0)
- Evans, R. (1984). *Strained mercy: The economics of Canadian health care*. Butterworths.
- Fehr, E., & Camerer, C. (2007). Social neuroeconomics: The neural circuitry of social preferences. *Trends in Cognitive Sciences*, 11(10), 419–427. <https://doi.org/10.1016/j.tics.2007.09.002>
- Feldstein, M. (1970). The rising price of physicians' services. *The Review of Economics and Statistics*, 52(2), 121–133. <https://doi.org/10.2307/1926113>
- Fiebig, D. G., Keane, M. P., Louviere, J., & Wasi, N. (2010). The generalized multinomial logit model: Accounting for scale and coefficient heterogeneity. *Marketing Science*, 29(3), 393–421. <https://doi.org/10.1287/mksc.1090.0508>
- Gaynor, M., & Town, R. (2012). Competition in health care markets. In P. M, M. T & B. PP (Eds.), *Handbook of health economics*. Elsevier.
- Ge, G., & Godager, G. (2021). Predicting strategic medical choices: An application of a quantal response equilibrium choice model. *Journal of Choice Modelling*, 39, 100282. <https://doi.org/10.1016/j.jocm.2021.100282>
- Godager, G., Iversen, T., & Ma, C.-T. A. (2015). Competition, gatekeeping, and health care access. *Journal of Health Economics*, 39, 159–170. <https://doi.org/10.1016/j.jhealeco.2014.11.005>
- Godager, G., & Wiesen, D. (2013). Profit or patients' health benefit? Exploring the heterogeneity in physician altruism. *Journal of Health Economics*, 32(6), 1105–1116. <https://doi.org/10.1016/j.jhealeco.2013.08.008>
- Gosden, T., Bowler, I., & Sutton, M. (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. <https://doi.org/10.1177/135581960000500404>
- Gravelle, H., Liu, D., Propper, C., & Santos, R. (2019). Spatial competition and quality: Evidence from the English family doctor market. *Journal of Health Economics*, 68, 102249. <https://doi.org/10.1016/j.jhealeco.2019.102249>
- Gravelle, H., Scott, A., Sivey, P., & Yong, J. (2016). Competition, prices and quality in the market for physician consultations. *The Journal of Industrial Economics*, 64(1), 135–169. <https://doi.org/10.1111/joie.12098>
- Greene, W. (2007). *NLOGIT version 4.0*. Referecne guide. Econometric Software Inc.
- Hanson, K., & Jack, W. (2010). Incentives could induce Ethiopian doctors and nurses to work in rural settings. *Health Affairs*, 29(8), 1452–1460. <https://doi.org/10.1377/hlthaff.2009.0164>
- Hennig-Schmidt, H., Selten, R., & Wiesen, D. (2011). How payment systems affect physicians' provision behaviour--An experimental investigation. *Journal of Health Economics*, 30(4), 637–646. <https://doi.org/10.1016/j.jhealeco.2011.05.001>
- Hennig-Schmidt, H., & Wiesen, D. (2014). Other-regarding behavior and motivation in health care provision: An experiment with medical and non-medical students. *Social Science & Medicine*, 108, 156–165. <https://doi.org/10.1016/j.socscimed.2014.03.001>

- Hess, S., & Train, K. (2017). Correlation and scale in mixed logit models. *Journal of Choice Modelling*, 23, 1–8. <https://doi.org/10.1016/j.jocm.2017.03.001>
- Hole, A., & Kolstad, J. (2010). *Mixed logit estimation of willingness to pay distributions: A comparison of models in preference and WTP space using data from a health-related choice experiment*. Working paper 03/10. Department of Economics, University of Bergen.
- Iversen, T., & Lurås, H. (2000). Economic motives and professional norms: The case of general medical practice. *Journal of Economic Behavior & Organization*, 43(4), 447–470. [https://doi.org/10.1016/s0167-2681\(00\)00130-x](https://doi.org/10.1016/s0167-2681(00)00130-x)
- Jack, W. (2005). Purchasing health care services from providers with unknown altruism. *Journal of Health Economics*, 24(1), 73–93. <https://doi.org/10.1016/j.jhealeco.2004.06.001>
- Jia, L., Meng, Q., Scott, A., Yuan, B., & Zhang, L. (2021). Payment methods for healthcare providers working in outpatient healthcare settings. *Cochrane Database of Systematic Reviews*, 2021(1). <https://doi.org/10.1002/14651858.cd011865.pub2>
- Johar, M. (2012). Do doctors charge high income patients more? *Economics Letters*, 117(3), 596–599. <https://doi.org/10.1016/j.econlet.2012.07.024>
- Johar, M., Jones, G., & Savage, E. (2014). What explains the quality and price of GP services? An investigation using linked survey and administrative data. *Health Economics*, 23(9), 1115–1133. <https://doi.org/10.1002/hec.3071>
- Joyce, C., Scott, A., Jeon, S.-H., 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), 50. <https://doi.org/10.1186/1472-6963-10-50>
- Kolstad, J. R. (2011). How to make rural jobs more attractive to health workers. Findings from a discrete choice experiment in Tanzania. *Health Economics*, 20(2), 196–211. <https://doi.org/10.1002/hec.1581>
- Kolstad, J. T. (2013). Information and quality when motivation is intrinsic: Evidence from Surgeon report cards. *The American Economic Review*, 103, 2875–2910. <https://doi.org/10.3386/w18804>
- Kolstad, J., & Lindkvist, I. (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. <https://doi.org/10.1093/heapol/czs063>
- Kruk, M., Johnson, J., Gyakobo, M., Peter Agyei-Baffour, P., Kwesi Asabir, K., Kotha, S., Kwansah, J., Nakua, E., Snow, R., & Dzodzomenyo, M. (2010). Rural practice preferences among medical students in Ghana: A discrete choice experiment. *Bulletin of the World Health Organization*, 88(5), 333–341. <https://doi.org/10.2471/blt.09.072892>
- Lagarde, M., & Blaauw, D. (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. <https://doi.org/10.1186/1478-4491-7-62>
- Lagarde, M., & Blaauw, D. (2014). Pro-social preferences and self-selection into jobs: Evidence from South African nurses. *Journal of Economic Behavior & Organization*, 107, 136–152. <https://doi.org/10.1016/j.jebo.2014.09.004>
- Lancsar, E., & Louviere, J. (2008). Conducting discrete choice experiments to inform healthcare decision making. *Pharmacoeconomics*, 26(8), 661–677. <https://doi.org/10.2165/00019053-200826080-00004>
- Li, J. (2018). Plastic surgery or primary care? Altruistic preferences and expected specialty choice of U.S. Medical students. *Journal of Health Economics*, 62, 45–59. <https://doi.org/10.1016/j.jhealeco.2018.09.005>
- Liu, T., & Ma, C.-T. A. (2013). Health insurance, treatment plan, and delegation to altruistic physician. *Journal of Economic Behavior & Organization*, 85, 79–96. <https://doi.org/10.1016/j.jebo.2012.11.002>
- Louviere, J. J., & Lancsar, E. (2009). Choice experiments in health: The good, the bad, the ugly and toward a brighter future. *Health Economics, Policy and Law*, 4, 527–546. <https://doi.org/10.1017/s1744133109990193>
- Louviere, J., Pihlens, D., & Carson, R. (2011). Design of discrete choice experiments: A discussion of issues that matter in future applied research. *Journal of Choice Modelling*, 4, 1–8. [https://doi.org/10.1016/s1755-5345\(13\)70016-2](https://doi.org/10.1016/s1755-5345(13)70016-2)
- Makris, M. (2009). Incentives for motivated agents under an administrative constraint. *Journal of Economic Behavior & Organization*, 71(2), 428–440. <https://doi.org/10.1016/j.jebo.2009.03.013>
- Markis, M., & Siciliani, L. (2013). Optimal incentive schemes for altruistic providers. *Journal of Public Economic Theory*, 15(5), 675–699. <https://doi.org/10.1111/jpet.12047>
- McGuire, T. (2000). Physician agency. In J. Newhouse & A. J. Culyer (Eds.), *Handbook of health economics*. Elsevier.
- McIsaac, M., Scott, A., & Kalb, G. (2015). The supply of general practitioners across local areas: Accounting for spatial heterogeneity. *BMC Health Services Research*, 15(1), 450. <https://doi.org/10.1186/s12913-015-1102-y>
- Mooney, G., & Ryan, M. (1993). Agency in health care: Getting beyond first principles. *Journal of Health Economics*, 12(2), 125–135. [https://doi.org/10.1016/0167-6296\(93\)90023-8](https://doi.org/10.1016/0167-6296(93)90023-8)
- Olivella, P., & Siciliani, L. (2017). Reputational concerns with altruistic providers. *Journal of Health Economics*, 55, 1–13. <https://doi.org/10.1016/j.jhealeco.2017.05.003>
- Quaife, M., Terris-Prestholt, F., Di Tanna, G. L., & Vickerman, P. (2018). How well do discrete choice experiments predict health choices? A systematic review and meta-analysis of external validity. *The European Journal of Health Economics*, 19(8), 1053–1066. <https://doi.org/10.1007/s10198-018-0954-6>
- Rizzo, J. A., & Zeckhauser, R. J. (2003). Reference incomes, loss aversion, and physician behavior. *The Review of Economics and Statistics*, 85(4), 909–922. <https://doi.org/10.1162/003465303772815817>
- Rizzo, J. A., & Zeckhauser, R. J. (2007). Pushing incomes to reference points: Why do male doctors earn more? *Journal of Economic Behavior & Organization*, 63(3), 514–536. <https://doi.org/10.1016/j.jebo.2005.06.007>
- Santos, R., Gravelle, H., & Propper, C. (2015). Does quality affect patients’ choice of doctor? Evidence from England. *The Economic Journal*, 127(600), 445–494. <https://doi.org/10.1111/econj.12282>

- Schaumans, C. (2015). Prescribing behavior of general practitioners: Competition matters. *Health Policy*, 119(4), 456–463. <https://doi.org/10.1016/j.healthpol.2014.12.018>
- Scott, A. (2001). Eliciting GPs' preferences for pecuniary and non-pecuniary job characteristics. *Journal of Health Economics*, 20(3), 329–347. [https://doi.org/10.1016/s0167-6296\(00\)00083-7](https://doi.org/10.1016/s0167-6296(00)00083-7)
- Scott, A., Witt, J., Humphreys, J., Joyce, C., Kalb, G., Jeon, S. H., & Mcgrail, M. (2013). Getting doctors into the bush: General Practitioners' preferences for rural location. *Social Science & Medicine*, 96, 33–44. <https://doi.org/10.1016/j.socscimed.2013.07.002>
- Serra, D., Serneels, P., & Barr, A. (2011). Intrinsic motivations and the non-profit health sector: Evidence from Ethiopia. *Personality and Individual Differences*, 51(3), 309–314. <https://doi.org/10.1016/j.paid.2010.04.018>
- Siciliani, L. (2009). Paying for performance and motivation crowding out. *Economics Letters*, 103(2), 68–71. <https://doi.org/10.1016/j.econlet.2009.01.022>
- Smith, R., Lagarde, M., Blaauw, D., Goodman, C., English, M., Mullei, K., Pagaiya, N., Tangcharoensathien, V., Erasmus, E., & Hanson, K. (2012). Appealing to altruism: An alternative strategy to address the health workforce crisis in developing countries? *Journal of Public Health*, 35(1), 164–170. <https://doi.org/10.1093/pubmed/fds066>
- Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.
- Ubach, C., Scott, A., French, F., Awramenko, M., & Needham, G. (2003). What do hospital consultants value about their jobs? A discrete choice experiment. *British Medical Journal*, 326(7404), 1432–1437. <https://doi.org/10.1136/bmj.326.7404.1432>
- Vujicic, M., Alfano, M., Shengelia, B., & Witter, S. (2010). *Attracting doctors and medical students to rural Vietnam: Insights from a discrete choice experiment*. HNP discussion papers. World Bank.
- Wang, J., Iversen, T., Hennig-Schmidt, H., & Godager, G. (2020). Are patient-regarding preferences stable? Evidence from a laboratory experiment with physicians and medical students from different countries. *European Economic Review*, 125, 103411. <https://doi.org/10.1016/j.euroecorev.2020.103411>
- Wordsworth, S., Skåtun, D., Scott, A., & French, F. (2004). Preferences for general practice jobs: A survey of principals and sessional GPs. *British Journal of General Practice*, 54, 740–746.
- Yong, J., Scott, A., Gravelle, H., Sivey, P., & Mcgrail, M. (2018). Do rural incentives payments affect entries and exits of general practitioners? *Social Science & Medicine*, 214, 197–205. <https://doi.org/10.1016/j.socscimed.2018.08.014>
- Zaresani, A., & Scott, A. (2021). Is the evidence on the effectiveness of pay for performance schemes in healthcare changing? Evidence from a meta-regression analysis. *BMC Health Services Research*, 21(1), 175. <https://doi.org/10.1186/s12913-021-06118-8>

How to cite this article: Scott, A., & Sivey, P. (2022). Motivation and competition in health care. *Health Economics*, 31(8), 1695–1712. <https://doi.org/10.1002/hec.4533>

APPENDIX 1

TABLE A1 GMNL models used to recover individual-specific marginal utility of income (MUY)

	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, < 5000	−0.225*** (0.024)	0.506*** (0.027)	−0.473*** (0.044)	0.930 (0.032)
Location: Coastal, < 5000	0.207*** (0.022)	0.376*** (0.030)	0.188*** (0.043)	0.679 (0.032)
Location: Town, 5000–20,000	−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)	

(Continues)

TABLE A1 (Continued)

	Forced choice model		Status-quo model	
	Coefficient mean	Coefficient SD	Coefficient mean	Coefficient SD
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			-3.200*** (0.061)	
Tau	0.211*** (0.000)		0.162*** (0.000)	
Gamma	0.551*** (0.033)		0.002 (0.013)	
Observations	27,852		26,661	
Individuals	3207		3018	
Log likelihood	-17488		-8953	
BIC	35,323		18,203	
AIC	35,043		5	
Model chi-sq	3635.9*** (34 df)		40,673.2*** (29df)	

Note: The coefficient of earnings is assumed to have a lognormal distribution, hence the estimated means and standard deviations of $\ln(\beta_{\text{earn}})$ are presented. Other coefficients with standard deviations are assumed to have normal distributions. Models are estimated using maximum simulated likelihood with 500 Halton draws in NLOGIT 5. For more detail on the methods of the DCE, see Scott et al. (2013).