



Choice of hospital: Which type of quality matters?

Nils Gutacker ^{a,*}, Luigi Siciliani ^b, Giuseppe Moscelli ^a, Hugh Gravelle ^a

^a Centre for Health Economics, University of York, York, United Kingdom

^b Department of Economics and Related Studies, University of York, York, United Kingdom



ARTICLE INFO

Article history:

Received 19 November 2015

Received in revised form 27 June 2016

Accepted 16 August 2016

Available online 22 August 2016

JEL classification:

I11

I18

Keywords:

Quality

Demand

Healthcare

Hospitals

Competition

Patient reported outcomes

ABSTRACT

The implications of hospital quality competition depend on what type of quality affects choice of hospital. Previous studies of quality and choice of hospitals have used crude measures of quality such as mortality and readmission rates rather than measures of the health gain from specific treatments. We estimate multinomial logit models of hospital choice by patients undergoing hip replacement surgery in the English NHS to test whether hospital demand responds to quality as measured by detailed patient reports of health before and after hip replacement. We find that a one standard deviation increase in average health gain increases demand by up to 10%. The more traditional measures of hospital quality are less important in determining hospital choice.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Many European healthcare systems have recently extended patients' right to choose their provider of elective hospital care (Vrangbaek et al., 2012). One of the aims is to encourage hospitals to compete for patients by improving quality (Besley and Ghatak, 2003). A prerequisite for this is that patients' choice of hospital is influenced by quality (Faber et al., 2009; Marshall et al., 2004). But quality is multi-dimensional and the implications of increased incentives to compete on quality depend on which aspects of hospital quality affect demand: if demand responds to a limited set of quality dimensions then hospital attempts to attract patients may be misdirected.

Studies of quality and choice of hospital have measured quality by mortality rates, either at hospital level or for specific conditions (e.g. Sivey, 2008; Beckert et al., 2012; Moscone et al., 2012; Gaynor et al., 2012; McConnell et al., 2016), readmission rates (McConnell et al., 2016; Moscone et al., 2012; Varkevisser et al., 2012), or hospital reputation and composite scores (Bundorf et al., 2009; Pope, 2009; Ruwaard and Douven, 2014; Varkevisser

et al., 2010, 2012).¹ See Brekke et al. (2014) for an overview. The studies typically report a positive association between quality measures and hospital demand.

But these quality measures, even those that are condition-specific, are incomplete, revealing little about the changes in health and functioning that patients will experience as the result of treatment (Appleby and Devlin, 2004). In this paper we use newly available data to examine whether choice of hospital is also influenced by quality measures based on detailed patient reports of health and functioning before and after treatment.

Since April 2009, all hospitals in England have been required to collect patient-reported outcome measures (PROMs) for National Health Service (NHS)-funded patients undergoing hip and knee replacement, varicose vein surgery or groin hernia repair (Department of Health, 2008).² PROMs are detailed validated questionnaires used to measure patients' health status, functioning, and health-related quality of life before and after treatment. Hospital quality measures derived from PROMs improve over 'failure' measures such as mortality or emergency readmission rates in several ways. First, they capture more dimensions of health and more sensitively (Appleby

* Corresponding author. Centre for Health Economics, University of York, Heslington, York YO10 5DD, United Kingdom. Fax: +44 1904 321454.

E-mail address: nils.gutacker@york.ac.uk (N. Gutacker).

¹ Early studies from the US include Folland (1983), Luft et al. (1990), Burns and Wholey (1992), Hodgkin (1996), Tay (2003), Howard (2005) and Ho (2006).

² Similar initiatives are underway or being considered in other health systems, including Sweden (Rolfson et al., 2011), Australia (Arthroplasty Clinical Outcomes Registry, 2015), Canada (McGrail et al., 2012), the Netherlands (Landelijke Registratie Orthopedische Implantaten, 2013) and the US (Basch et al., 2013).

and Devlin, 2004; Gutacker et al., 2013). Second, they embody much better adjustment for case-mix, because they incorporate the rich pre-operative PROMs data. Finally, PROMs reflect the patients' views on their health and health improvement. Hospital-specific average risk-adjusted changes in patient health status have been produced and disseminated online to support the Department of Health's ambition that they would be used "by patients and [general practitioners (GPs)] exercising choice" (Department of Health, 2008, p. 6).^{3,4}

We use newly available PROMs-based quality data to estimate a multinomial logit model of patient choice of hospital for hip replacement surgery to test whether hospital demand responds to the aspects of quality captured by PROMs, as well as to more conventional measures such as mortality and readmission rates.⁵ Our main empirical strategy relies on the use of lagged hospital quality variables to identify the effect of quality on hospital choice: patients are assumed to use quality information that is made publicly available with a one-year lag and that can therefore not be affected by current choices. We test the robustness of our results by estimating models with hospital fixed effects, and by using closely matched emergency hip replacement patients as a control group.

We find that patients are more likely to choose providers producing greater health gains. In our baseline specification, the elasticity of hospital demand with respect to the average health gain (measured in terms of the Oxford Hip Score) is 1.4, whereas the demand elasticities with respect to the readmission rate and the mortality rate are -0.02 and -0.004 . A one standard deviation increase in the health gain increases demand by up to 9.8% whereas one standard deviation increases in readmission and mortality rates reduce demand by 6.8% and 0.7%, respectively. This suggests that hospitals' abilities to improve their patients' health reflect an important dimension of quality that is valued by patients and not captured by other quality measures. Our robustness analyses, both with emergency patients as a control group and using hospital fixed effects, also find statistically significant, though smaller, marginal willingness to travel for higher PROM quality and lower elasticities of demand. For example, using emergency patients as a control group we find the demand elasticity to PROM quality to be 0.6 and a one standard deviation increase in health gain is estimated to increase demand by 4.4%.

Our study makes a number of contributions. It is the first to explore whether hospital demand responds to quality as measured by average health gains derived from patient self-reported outcome measures. Second, we make novel use of pre-operative individual level PROMs data to explore whether sicker patients travel

farther and choose hospitals with higher quality of care. Such self-selection can bias performance assessments ('distance bias') and create perverse incentives in the context of e.g. pay-for-performance schemes (Etzioni et al., 2013; Gowrisankaran and Town, 1999). Previous studies have either relied on instrumental variable approaches to allow for the effect of unobserved pre-operative health status on demand (Geweke et al., 2003; Gowrisankaran and Town, 1999) or have used simple measures of comorbidity and past utilisation. Our data have rich information on pre-operative health, which allows us to address this question directly. Third, our paper contributes to the small literature on hospital choice in publicly funded health systems where prices are regulated and demand is rationed by waiting time (Beckert et al., 2012; Gaynor et al., 2012; Moscone et al., 2012; Sivey, 2012). Our analysis differs from Beckert et al. (2012), who also study choice of provider for hip replacement surgery in England, in that we use provider quality measures that are procedure-specific and more directly related to the quality of care provided⁶, explore the role of pre-operative health status, and model the entire NHS market, including private providers of NHS-funded care.

More broadly, the analysis relates to the literature on hospital report cards, which has mostly focused on the publication of coronary artery bypass graft (CABG) surgery quality scores in a number of US states. For example, Chou et al. (2014) investigate whether information on quality gives stronger incentives to improve care for hospitals that face more competition. They find that after the online publication of CABG report cards in Pennsylvania, hospitals in more competitive areas used more resources per patient and had lower mortality for more severe patients than those in less competitive areas. Wang et al. (2011) find that the probability of patients receiving CABG surgery from low-performing surgeons was significantly reduced following the publication of report cards. In contrast, Epstein (2010) failed to detect changes in referrals following the introduction of report cards in Pennsylvania compared to the counterfactual state Florida, and this may be due to referring physicians already having expectations about the relative performance of cardiac surgeons. The result is consistent with Dranove and Sfekas (2008) who find that the introduction of report cards in New York State was less effective when the ratings confirmed prior beliefs about quality. Dranove et al. (2003) found that the introduction of report cards led to a reduction in average severity with doctors trying to avoid high-risk patients. Cutler et al. (2004) observed improvements in quality for New York State hospitals that received a negative report card, suggesting that healthcare providers are responsive to quality reporting. This literature analyses either aggregate demand patterns or individual choice behaviour using discrete choice models (similar to the approach adopted in this study) and takes quasi-experimental approaches or panel data techniques to establish causality.

In the next section we describe the data and in Section 3 we set out the econometric model and our strategy to mitigate potential endogeneity bias. In Section 4 we present the estimated marginal utilities of different aspects of quality and other hospital characteristics and show how these vary with observed patient characteristics. Section 5 reports estimates of the effect of providers' quality on their own demand and on demand of their competitors. In Section 6 we discuss the key results.

³ The English NHS is funded by general taxation and there are no charges to patients (except for a small charge for dispensed medicines). Providers receive a nationally set prospective price for hip replacement paid for by the patient's Primary Care Trust from its capitated budget. There is a list system in general practice and patients wishing to access elective care must be referred to a hospital by their GP. If a hospital specialist decides the patient should be treated she is placed on a waiting list. Patients have the right to be treated at any qualified provider (NHS or private). Providers must accept all clinically appropriate referrals from GPs and cannot close their waiting lists to new patients.

⁴ Some patients might access this information directly, whereas others might rely on their GPs, who act as their agent, to retrieve, interpret and communicate this information. In some instances, patients may not even be willing or able to make a choice and their referring GP may choose the most appropriate hospital for them, i.e. the GP acts as an agent to the patient. It is generally not possible to distinguish between decision makers using administrative data. For simplicity, we will henceforward denote the decision-maker as the patient.

⁵ Information on the average change in PROM scores by hospital, including the change in the Oxford Hip Score as used in our analysis, is reported to the general public through the NHS Choices website (<http://www.nhs.uk/>) and the Health & Social Care Information Centre website (<http://www.hscic.gov.uk/proms>). The same public sources also provide information on procedure-specific mortality, emergency readmission and revision rates.

⁶ Beckert et al. (2012) model hospital quality using hospital-wide mortality and MRSA infection rates. Aggregate hospital level quality indicators, such as the summary hospital mortality indicator (SHMI) used in the English NHS, do not correlate well with procedure-specific outcome measures (Gravelle et al., 2014). In 2010/11, the Pearson correlation coefficients between SHMI and the quality measures used in this study were -0.09 (change in OHS), -0.07 (emergency readmission rate), -0.005 (revision rate) and 0.08 (mortality rate), respectively.

2. Data

We use patient-level data from Hospital Episode Statistics (HES) for all elective admissions for patients aged 18 or over who underwent NHS-funded primary (i.e. non-revision) hip replacement surgery between April 2010 and March 2013 in NHS or private providers.⁷ HES contains rich information on patients' demographic and medical characteristics, small area of residence and on the hospital stay. Privately funded patients are excluded from our analysis.⁸ We also exclude patients attending providers with less than 30 patients in that year.

We derive four patient variables from HES: age, gender, the number of emergency admissions during the 365 days prior to their hip replacement admission, and the number of Elixhauser comorbid conditions recorded in admissions in the previous year (Elixhauser et al., 1998; Gutacker et al., 2015a). These are available for all patients. As a measure of income deprivation we use the 2010 Index of Multiple Deprivation (McLennan et al., 2011) to attribute to each patient the proportion of residents in households claiming means-tested benefits in their Lower Super Output Area (LSOA).⁹ We measure a patient's distance from a hospital as the straight-line distance from the centroid of their LSOA.¹⁰

All NHS-funded hip replacement patients are invited to report their health status and health-related quality of life before, and six months after, surgery using a paper-based questionnaire. The pre-operative PROM questionnaire is administered by the hospital either as part of the admission process or during the last outpatient appointment preceding admission. The post-operative questionnaire is administered by a central agency and posted to the patient. Participation in the PROM survey is compulsory for providers but optional for patients. Approximately 60% of patients provide complete pre- and postoperative PROM questionnaires that can be linked to their HES record (Gutacker et al., 2015b).

Each PROM questionnaire contains three instruments: the Oxford Hip Score (OHS), the EuroQoL-5D (EQ-5D) descriptive system, and the EuroQoL Visual Analogue Scale (EQ-VAS). The OHS is a condition-specific instrument with 12 questionnaire items regarding hip-related functioning and pain (Dawson et al., 1996). Each item is scored on a five-point scale, with four indicating no problems and zero indicating severe problems. The overall score is calculated as the sum of all items and ranges from zero (worst) to 48 (best). Both EuroQoL instruments are generic PROMs, i.e. they can be applied to different health conditions (Brooks, 1996). Previous analysis showed substantial correlation between the EQ-5D and OHS (Neuburger et al., 2013). We focus on analysis with the OHS since it is a condition-specific measure and hence plausibly more likely to affect hospital choice for hip replacements. We report a choice model using the EQ-5D quality measures as sensitivity check.

⁷ See Department of Health (2008) for procedure codes. We exclude patients who had revision surgery. This ensures a more homogeneous sample and they are likely to return to the hospital providing the initial hip replacement.

⁸ Approximately 16% of hip replacement surgeries are funded privately, either out-of-pocket or through private insurance (Hunt et al., 2013). Privately funded patients treated in the private sector are not included in HES and we do not have information on them. We also exclude the small number (<2%) of patients treated in NHS hospitals who are privately funded.

⁹ LSOAs have a mean population of 1500 and are constructed to be homogeneous with respect to tenure and accommodation type.

¹⁰ Some NHS trusts perform hip replacement surgery across multiple sites, but the site codes for NHS hospitals are often poorly recorded in HES. Also, quality information for NHS providers is only recorded at trust level. We therefore locate hospitals using the postcode of the trust headquarter (for NHS), which is consistently recorded, or the hospital site (for ISTCs). Analysis of complete data reveals that multi-site operations are rare: approximately 80% of NHS trusts perform more than 80% of their surgery in a single site.

We use PROMs data in two ways. First, we use the risk-adjusted hospital-specific OHS change scores computed and published by the Health & Social Care Information Centre (HSCIC) as a measure of hospital quality (Health & Social Care Information Centre, 2013b). Data are reported by financial years (April to March). The HSCIC case-mix adjustment methodology takes into account a range of patient characteristics including age, sex, pre-operative PROM score, socio-economic status, comorbidity burden, whether the patient lives alone as well as other indicators of disability (Department of Health, 2012). The hospital-specific mean scores have been found to be robust to missing data (Gomes et al., 2015). They are also unlikely to be affected by survivorship bias since less than 0.1% of patients die during the six-month follow-up. Second, in some of our models, we use the information in the individual patients' pre-operative PROMs questionnaires to investigate whether their choice of provider is affected by pre-operative morbidity.

We also calculate risk-adjusted hospital-specific annual 28-day emergency readmission, 28-day mortality rates and 1-year revision rates after primary hip replacement surgery as additional quality measures. We link our HES data to the Office of National Statistics death records and apply the HSCIC case-mix adjustment as set out in the readmission outcome indicator specification (Health & Social Care Information Centre, 2013a).^{11,12} Quality metrics are not calculated for hospitals that treat less than 30 elective hip replacement patients in that year because they are considered too noisy for analysis.

We group providers into seven categories used by the National Patient Safety Agency (n.d.): NHS small/medium/large non-teaching trust, NHS teaching trust, NHS specialised orthopaedic provider, NHS multi-service provider, and NHS Primary Care Trusts (PCTs).¹³ We also distinguish NHS hospitals from privately operated Independent Sector Treatment Centres (ISTCs) that treat NHS patients. The size categorisation of NHS providers is based on the overall number of beds available and providers' annual revenue. Teaching trusts receive additional funding from the English Department of Health to train doctors and are more research active. Multi-service providers offer acute services as well as mental health and/or community services.¹⁴ Orthopaedic specialist hospitals only treat musculoskeletal conditions, and, like ISTCs, do not provide emergency care.

Finally, we derive from HES the proportion of patients in each hospital who waited longer than 120 days between the specialist's decision to add the patient to the waiting list and admission to hospital for treatment (the inpatient wait). Patients dislike waiting since health benefits from treatment are postponed and they are in pain and less mobile while waiting. Waiting times have been shown to reduce the overall demand for NHS elective care (Gravelle et al., 2002; Martin and Smith, 1999, 2003) and to influence choice of provider (Beckert et al., 2012; Gaynor et al., 2012; Moscone et al., 2012; Sivey, 2012). We focus on long waits because this is how waiting time targets for NHS providers are defined and thus reported in the press if hospitals miss targets. We also conduct sensitivity analysis using the median waiting time (in months) at each provider.

¹¹ This adjustment takes into account age (in 5-year bands), sex, socio-economic status, comorbidity burden as captured by the Charlson index and the number of emergency admissions in the last year.

¹² One hospital-year observation was dropped from the analysis because information on discharge date was missing for all patients.

¹³ During our study period, PCTs were responsible for purchasing care for their resident population and, with the exception of the Isle of Wight PCT, did not provide care themselves.

¹⁴ Such sites must provide at least two differing service functions, each with gross internal floor area representing more than 20% of the total gross internal site floor area for the whole site (Health & Social Care Information Centre, n.d.).

3. Methods

3.1. Model specification

We use a random utility choice model (McFadden, 1974). In our baseline specification utility of patient $i = 1, \dots, N$ at provider $j = 1, \dots, J$ at time $t = 1, \dots, T$ is:

$$U_{ijt} = V_{ijt} + v_{ijt} \\ = D'_{ij}\beta_{d,i} + D''_{ij}\beta_{d^2,i} + D'''_{ij}\beta_{d^3,i} + Q'_{jt-1}\beta_{q,i} + W'_{jt-1}\beta_{w,i} + Z'_j\beta_{z,i} + v_{ijt} \quad (1)$$

where V_{ijt} depends on the distance from the patient's residence to the hospital D_{ij} , hospital quality metrics Q_{jt-1} , waiting time W_{jt-1} , and a vector of time-invariant hospital characteristics Z_j . The error term v_{ijt} subsumes unobserved hospital characteristics and random utility. Patients choose from a set of hospitals M_{it} . Assuming v_{ijt} is iid extreme value yields the *multinomial logit* (MNL) model in which the probability that patient i chooses hospital j is:

$$P_{ijt} = \frac{\exp(V_{ijt})}{\sum_{j' \in M_{it}} \exp(V_{ijt'})} \quad (2)$$

We allow a non-linear effect of distance on utility by modelling linear, squared and cubic distance terms.¹⁵ We assume that anticipated utility at a provider is based on its previous period's quality and waiting time because relevant information are available only with a lag (see section 3.2). Varkevisser et al. (2012) make a similar assumption. In sensitivity checks, we also estimate models with contemporaneous waiting time and quality scores.

We allow preferences to vary across patients according to their observed characteristics. Thus the marginal utility of quality for patient i is:

$$\beta_{q,i} = \beta_q + X'_i\delta_q \quad (3)$$

and similar for distance, waiting time, and other hospital characteristics. All continuous covariates in X_i are mean centred and base categories for categorical characteristics are set to their mode. Thus, the vectors of coefficients β_d , β_{d^2} , β_{d^3} , β_q , β_w , β_z reflect the preferences of the *reference patient*.

We also estimate models which allow for unobserved patient heterogeneity in tastes over quality, with:

$$\beta_{q,i} = \beta_q + X'_i\delta_q + \sigma_q\alpha_i \quad (4)$$

where σ_q is the standard deviation of a zero mean normal variable and α_i is an unobserved patient effect. The latter may capture, for example, differences in the ability to access and interpret quality information. This *random coefficient multinomial logit* (RCMNL) or mixed logit model (Hensher and Greene, 2003; Train, 2003), unlike the MNL model, allows for unrestricted substitution patterns, thereby relaxing the assumption of independence of irrelevant alternatives (IIA). If $\sigma_q = 0$ then the RCMNL model reduces to the MNL model in (1).

While the MNL model has a closed form solution that can be estimated via maximum likelihood, the RCMNL needs to be approximated through simulation. To reduce the computational burden¹⁶ we assume uncorrelated normally distributed random coefficients for the quality metrics in Q_{jt-1} and no random coefficients

for other covariates. The RCMNL model is estimated with maximum simulated likelihood using 50 Halton draws.

All models are estimated in Stata 13 with *clogit* and the user-written command *mixlogit* (Hole, 2007b). Standard errors are clustered at the GP practice level to allow for GP-induced correlation across choices by patients in a practice.

3.2. Endogeneity

A necessary condition for β_q to be an unbiased estimate of the marginal utility of hospital quality (up to a linear transformation) is that the error term v_{ijt} is uncorrelated with any of the independent variables, i.e. all observed variables are exogenous. This condition may not hold for four reasons (Brekke et al., 2014; Gaynor et al., 2012; Varkevisser et al., 2012).

First, hospitals may learn by doing so that higher volume providers have higher quality (Gaynor et al., 2005; Luft et al., 1987). Thus, changes in demand could affect quality and induce simultaneity bias. A pre-PROM study with English 1997–2002 hip replacement data found that 30-day in-hospital mortality was higher in low volume hospitals that treated less than 100 patients per year but found no volume effect above this threshold (Judge et al., 2006). The average NHS hospital in our sample treated over 361 NHS patients per year and the average ISTC treated 155 NHS patients per year as well as an unknown number of private patients. A more recent study using the OHS reported no relationship between hospital volume and quality (Varagunam et al., 2015). Furthermore, since demand is likely to respond to past, rather than current, quality, our modelling strategy prevents such bias: demand changes in period t cannot affect quality in $t-1$. We therefore believe that simultaneity arising from the effect of volume on quality is not a problem with our data and modelling approach.

Second, because of short run capacity constraints, changes in demand will also affect waiting time in the same period (Gaynor et al., 2012).¹⁷ While our primary interest is not in the effect of waiting time on demand, we are concerned that any bias introduced through endogenous variables will filter through to our estimate of β_q (Wooldridge, 2002). Again, the use of lagged waiting time measures guards against this type of bias.

Third, sicker patients may choose higher quality hospitals or hospitals may turn away or discourage patients with characteristics that make them less likely to achieve a large improvement in health status. If such systematic selection occurs and is not controlled for in the calculation of hospital quality scores then those scores would in part be determined by patients' choices or provider selection. However, provider quality scores are adjusted for a rich set of demographic, socio-economic, and morbidity patient characteristics, including, in the case of PROMs, the patients' self-reported pre-operative health status. Hence, we do not believe that unobserved patient selection is likely to bias the quality scores significantly.

Finally, the error term in (1) may be the sum of unobserved hospital characteristics ξ_{jt} and iid random patient utility, i.e. $v_{ijt} = \xi_{jt} + \epsilon_{ijt}$ where ξ_{jt} affect demand and are correlated with observed covariates (Ho, 2006; Jung et al., 2011). For example, hospitals in areas with better amenities may attract better staff thereby ensuring higher observed clinical quality but also unobserved interpersonal aspects of quality. Our assumption that patients use information on previous period quality and waiting times when choosing hospitals does not remove omitted variable bias operating through unobserved non-transitory hospital characteristics. We therefore undertake two types

¹⁵ The results are noticeably different from a model including distance and distance² only, but are very similar to one that also included the fourth power.

¹⁶ Even after imposing those constraints the RCMNL model with our baseline specification still took over 3 weeks to compute on a high-performance computing system.

¹⁷ It may also be that supply and demand are determined simultaneously, i.e. hospitals react to demand shocks by adjusting their supply, e.g. by performing more surgeries on weekends. We do not consider this in our model explicitly, although the use of lagged waiting time circumvents this problem as well.

of sensitivity analyses to explore the size of the potential omitted variable bias. Our first approach is to estimate the choice model in (1) with alternative-specific time-invariant fixed effects (FEs) (Ho, 2006; Hodgkin, 1996; Monstad et al., 2006; Sivey, 2012). These hospital FEs capture the utility of non-transitory unobserved hospital characteristics. The coefficients on observed hospital characteristics are now identified solely through variation within providers over time, thereby removing any endogeneity bias operating through unobserved time-invariant characteristics. However, since hospital quality varies little over time (see Section 4.1) and we only observe providers over three years this approach may yield imprecise estimates of the marginal utility of quality. Also, because the market structure changes over time, due to the opening of new independent sector treatment centres, the FEs do not correspond to observed market shares in each time period. This may bias estimates if incumbent providers differ systematically from new entries. We therefore also estimate a model based on NHS trusts only, whose numbers are relatively stable over time.

Our second approach is to follow Pope (2009) (see also Moscone et al., 2012 and Gaynor et al., 2012) and gauge the possible impact of unobserved hospital heterogeneity by using a control group of emergency hip replacement patients whose choice of provider is less responsive to quality and waiting time. The majority of emergency hip replacement patients suffer from a fractured neck of the femur as a result of a fall. The key idea is that in an emergency patients do not have time to compare the quality of hospitals (or at least not as easily as for a planned procedure). The official recommendation is that hip fracture patients should be treated within 48 hours because further delays are linked to worse outcomes (NICE, 2011; Moja et al., 2012).¹⁸ Conversely, elective patients have time to gather information, consult websites, family doctors, etc. We therefore expect elective patients to respond more to quality than emergency patients, whose scope to choose providers based on quality is clearly limited given the time-sensitive nature of their condition.

Emergency hip replacement surgery is undertaken by the same orthopaedic departments that provide elective surgery and so measures of the quality of elective care also apply. However, given the relatively urgent nature of their condition, we expect provider choice by emergency hip replacement patients to be largely unaffected by publicly reported information on quality. To the extent that they exercise choice they are likely to be influenced by distance to providers and time-invariant unobserved factors that also matter to elective patients, such as long-standing reputation or dimensions of accessibility not captured by our distance measure (e.g. parking charges or connection to the public transport system).

If we assume that emergency patients' demand is entirely inelastic to observed quality and that they do not wait,¹⁹ but value the same unobserved hospital characteristics as elective patients, then their true utility is given by:

$$U_{ijt}^{Emer} = D'_{ij} \beta_{d,i}^{Emer} + D'_{ij} \beta_{d^2,i}^{Emer} + D'_{ij} \beta_{d^3,i}^{Emer} + v_{ijt} \quad (5)$$

If we estimate the model specified in (1) for emergency patients and find $\hat{\beta}_q^{Emer} \neq 0$, we conclude that $cov(Q_{jt-1}, \xi_{jt}) \neq 0$. If we also assume that elective and emergency patients have the same preferences for unobserved hospital characteristics, then the effect

of quality on elective demand, purged of omitted variable bias, is $\beta_q^A = \beta_q^{Elec} - \beta_q^{Emer}$. Hence, if these assumptions are valid, the strategy is similar to a control group design. Since coefficients in separate MNL models may be scaled differently, we estimate a pooled model for elective and emergency patients by interacting all covariates with an indicator variable for emergency. This forces the scaling to be the same. The coefficients on the interaction terms are estimates of β_k^A for $k \in [d, d^2, d^3, q, w, z]$.

If emergency patients choose using information on the same type of quality as elective patients,²⁰ or care about emergency quality that correlates with it, or if unobserved hospital characteristics have different effects on choices by emergency and elective patients and are correlated with observed quality, then β_q^A can no longer be interpreted as an unbiased estimate of the effect of quality on elective demand. If unobserved hospital factors are not correlated with quality, then β_k^A reflects the differences in preferences between elective and emergency patients. In this case, we expect that $\beta_q^A > 0$: elective patients will be more sensitive to quality than emergency patients, at least in part because they have more time and can exert more effort to identify high quality care providers.

Finally, we note that our findings may be of interest even without a strict causal interpretation since they reveal whether patient demand patterns tend to favour high-quality hospitals (e.g. hospitals with a good reputation may invest more in quality because of intrinsic motivation).

3.3. Elasticities, changes in demand and willingness to travel

The estimated coefficients on hospital characteristics are estimates of the marginal utilities. Since the utility function is unique only up to a linear transformation, the coefficients only convey information about the sign of marginal utility of hospital characteristics and hence about the sign of the effect of quality on demand. The ratio of marginal utilities (the negative of the marginal rate of substitution) is unaffected by linear transformations and so provides quantitative information about patient preferences. We estimate the reference patient's willingness to travel (WTT) for a one standard deviation (SD) increase in quality as:

$$\begin{aligned} WTT &= \frac{\partial D_{ij}}{\partial Q_j} \bigg|_{U_{ij}} SD(Q) = - \frac{\partial U_{ij}}{\partial Q_j} / \frac{\partial U_{ij}}{\partial D_j} SD(Q) \\ &= \frac{-\beta_q}{\beta_d + 2\beta_{d^2}D + 3\beta_{d^3}D^2} SD(Q) \end{aligned} \quad (6)$$

where D is the average distance to elective patients' chosen hospitals. WTT is the extra distance in kilometres that the reference patient located the average distance away from a provider would be willing to travel to that provider if its quality was increased by $SD(Q)$, where $SD(Q)$ is averaged across hospitals and years. To allow comparisons of WTTs across models based on different sample definitions we evaluate all WTT at the values of D and $SD(Q)$ for the sample used to estimate our baseline specification (see Table 2). We estimate standard errors by the delta method (Hole, 2007a).

We are also interested in whether providers could attract more patients by improving their quality. Expected demand at provider j is $Y_{jt} = \sum_{i \in S_{jt}} P_{ijt}$, where S_{jt} is the set of patients whose choice set includes provider j , i.e. for whom $j \in M_{it}$. Following Santos et al. (2015) we calculate the average partial effect of a one SD increase in quality on provider j 's demand, i.e. demand responsiveness to quality, as:

¹⁸ The NHS choices website suggests that "if you think you've fractured your hip, you will need to go to the hospital as soon as possible. Dial 999 to request an ambulance. Try not to move while you're waiting for the ambulance and make sure you keep warm". The text gives an idea of the criticality of the condition.

¹⁹ Emergency patients are usually treated within hours of their hip fracture, not weeks or months. Analysis of our data revealed that elective waiting time is only weakly correlated with the volume of emergency patients, suggesting that there is separate supply for those two types of patients.

²⁰ As with elective patients, we do not observe who chooses the hospital for emergency hip replacement. This may be the patient, a family member, GP, or the ambulance crew.

Table 1
Descriptive statistics – elective sample.

Variable	Obs	Mean	SD	ICC
Patient characteristics				
Distance travelled (in km)	171,472	14.6	17.4	
Distance travelled past closest provider (in km)	171,472	5.0	14.4	
Number of providers within 10 km radius	171,472	1.3	1.5	
Number of providers within 30 km radius	171,472	6.7	5.7	
Age	171,472	68.0	11.5	
Male	171,472	0.40	0.49	
Past utilisation	171,472	0.13	0.49	
Number of Elixhauser conditions	171,472	0.43	0.94	
Income deprivation	171,472	0.12	0.09	
Pre-operative Oxford Hip Score	114,187	17.4	8.2	
Provider characteristics				
Observed volume	568	305.4	209.1	94.7%
Waiting time (proportion waiting >120 days)	568	0.2	0.2	59.5%
Change in Oxford Hip Score	568	19.8	1.4	57.2%
28-day emergency readmission rate (%)	568	5.67	2.40	36.8%
1-year revision rate (%)	568	0.94	0.82	17.8%
28-day mortality rate (%)	568	0.17	0.36	3.5%

Obs = Observations; SD = Standard deviation; ICC = Intraclass correlation coefficient. Notes: Patient characteristics for patients choosing provider between April 2010 and March 2013. Waiting time and quality metrics are for financial years 2009/10 to 2011/12. Provider characteristics are unweighted. The intraclass correlation coefficient (ICC) measures the proportion of variation that occurs between providers, rather than over time.

$$\frac{\partial Y_{jt}}{\partial Q_{jt-1}} SD(Q) = SD(Q) \sum_{i \in S_{jt}} \frac{\partial P_{ijt}}{\partial Q_{jt-1}} = SD(Q) \sum_{i \in S_{jt}} \beta_q P_{ijt} (1 - P_{ijt}) \quad (7)$$

We report the mean of (7) over all providers and years.

We calculate the elasticity of demand of provider j with respect to own quality as:

$$E_{jt}^{Q_{jt-1}} = \sum_{i \in S_{jt}} \frac{\partial P_{ijt}}{\partial Q_{jt-1}} \frac{Q_{jt-1}}{\sum_{i \in S_{jt}} P_{ijt}} = \sum_{i \in S_{jt}} \beta_q P_{ijt} (1 - P_{ijt}) \frac{Q_{jt-1}}{\sum_{i \in S_{jt}} P_{ijt}} \quad (8)$$

We report the mean of (8), weighted by providers' predicted demand $\sum_{i \in S_{jt}} P_{ijt}$.

Finally, we compute the cross-elasticity of demand for provider j with respect to the quality of provider j' as:

$$E_{jt}^{Q_{j'}} = \sum_{i \in S_{jt} \cap S_{j'}} \frac{\partial P_{ijt}}{\partial Q_{j't-1}} \frac{Q_{j't-1}}{\sum_{i \in S_{jt}} P_{ijt}} = - \sum_{i \in S_{jt} \cap S_{j'}} \beta_q P_{ijt} P_{ij't} \frac{Q_{j't-1}}{\sum_{i \in S_{jt}} P_{ijt}} \quad (9)$$

with $j \neq j'$. Note that for some combinations of j and j' the cross-elasticity is zero because no patients have both providers in their choice sets.

4. Results

4.1. Descriptive statistics

Our main sample is 171,472 elective hip replacement patients treated in 228 providers during the period April 2010 to March 2013.²¹ Their average age is 68 years and 40% are male (Table 1). The average pre-operative OHS is 17.4 and 9.5% of patients have been admitted to hospital as an emergency at least once during the preceding 365 days (average number of admissions = 0.13). Self-reported

²¹ The number of providers varied over this period because of mergers, changes in coding and market entry, especially with respect to private facilities. There were 155 providers in 2010/11, 202 in 2011/12, and 211 in 2012/13, of which 16 (11.5%) in 2010/11, 62 (30.7%) in 2011/12, and 77 (36.5%) in 2012/13 were privately operated.

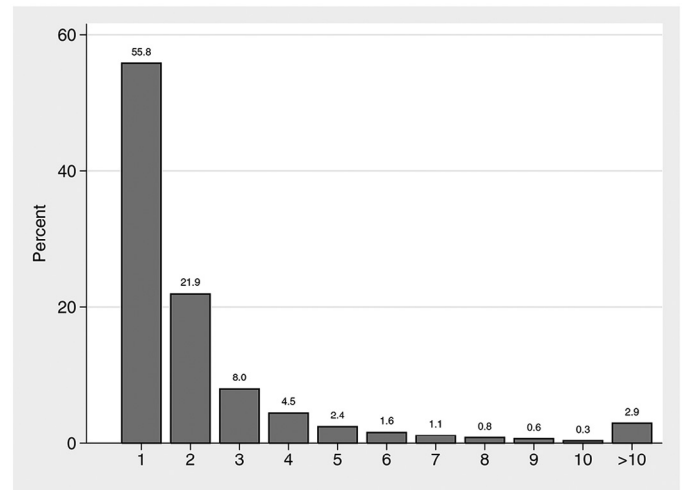


Fig. 1. Percentage of elective patients who went to their Nth nearest hospital.

pre-operative OHS is only weakly correlated with past emergency utilisation ($\rho = -0.10$) and the number of comorbidities ($\rho = -0.14$). This suggests that past emergency utilisation and comorbidity burden are poor proxies for current health status²² as experienced by the patient.

On average, patients have a choice of 7 providers within 30 km, with over 75% of patients having access to at least two different providers. Within 10 km there are on average 1.3 hospitals and over 14% of patients can choose between two or more providers. Conversely, 36% of patients have no options within a 10 km radius. To reduce computational burden we restrict patient choice sets to the 50 nearest providers. The 556 patients (0.3% of the sample) who chose a provider outside this set were dropped from the analysis.

Patients live on average 14.6 km from their chosen hospital. Fig. 1 shows that just under half (44.2%) of patients bypassed the local hospital and nearly a tenth (9.9%) bypassed the nearest three hospitals. On average, patients travel 5.0 km beyond their nearest hospital to be treated.

The hospital waiting time and quality scores are lagged by one year and are for financial years 2009/10 to 2011/12. There are much larger coefficients of variation for hospital emergency re-admission, revision and mortality rates than for OHS change scores. The provider OHS change scores are only weakly correlated with our waiting time measure ($\rho = -0.23$), readmission rates ($\rho = -0.28$), revision rates ($\rho = -0.07$) and mortality rates ($\rho = -0.05$). This suggests that choice models that are restricted to failure rates may not even indirectly pick up the effect of PROM measures on demand.

The intra-class correlation coefficient (ICC) shows that just over half of the observed variation in OHS (ICC = 57%) and our waiting time measure (ICC = 59%) is between providers rather than over time. This is less pronounced for readmission, revision and mortality rates (all ICC < 40%). The limited within (i.e. over time) variation in provider characteristics impedes the identification of associated coefficients in the fixed effects model.

4.2. Regression results

4.2.1. Main effects

The results for the RCMNL model (see Appendix Table A1) are very similar to the MNL model and the Hausman tests also did not

²² The correlations between these measures and the EQ-5D utility score are similar: $\rho = -0.10$ for past utilisation, and $\rho = -0.13$ for comorbidity burden.

reject the IIA assumption. We therefore concentrate on the MNL models reported in Tables 2–6.

Table 2 is our baseline specification with distance, lagged waiting time, the four lagged quality metrics, and indicators for the type of provider as well as interactions with patient age, gender, past utilisation, comorbidity, and local area income deprivation (we explore interactions with pre-operative OHS in section 4.2.2). The main effects are the estimated marginal utilities for the reference patient with mean or modal characteristics. The reference patient prefers shorter distances with the marginal disutility from distance declining with distance. She prefers providers with lower proportions of long-waits and specialised providers to non-specialised providers. She is also more likely to choose a public provider over a private provider after accounting for distance, waiting time and quality.²³

Reference patient demand is increasing with the OHS change score and decreasing with emergency readmission rates. The estimated WTT for a one SD increase in OHS is 0.9 km or 6% of the average distance travelled to the chosen provider. The WTT for a SD decrease in emergency readmission rates is 0.6 km. There is no statistically significant effect of procedure-specific mortality or revision rates on demand.

Using contemporaneous, rather than lagged, waiting time and quality measures, we find that the sign of all coefficients remain the same, but the coefficients on revision and mortality rates are now statistically significant (Appendix Table A2, model 1), possibly because of endogeneity in the contemporaneous quality measures. When we use the lagged median waiting time as a waiting time measure its coefficient is positive but not statistically significant (Appendix Table A2, model 2) and the other coefficients are unaffected. This lack of an effect of median waiting time may be explained by its historically low level, with patients being more concerned about excessive rather than average waits.²⁴ The HSCIC also produces hospital quality scores based on the case-mix adjusted change in the EQ-5D utility score. This is highly correlated with the OHS change score (Neuburger et al., 2013) and when we estimate the baseline specification with EQ-5D substituted for OHS we find similar WTT of 0.9 km for a one SD increase in the PROM measure (Appendix Table A2, model 3). We also test for a non-linear relationship between marginal utility and PROM quality by re-estimating the model with indicators for quartile groups of the OHS change score. We find a positive dose–response relationship: marginal utility increases proportionally between the first and second quartile, and between the second and third quartile, with no further increase between the third and fourth quartiles (available on request). Excluding independent sector treatment centres from patient choice sets (Appendix Table A2, model 4) has little impact, though the coefficient on revision rates becomes negative and statistically significant. Finally, results are robust to differing definitions of choice sets (i.e. 30 or 100 closest providers) (Appendix Table A3).

Very few healthcare systems collect and report information on health gains from surgery disaggregated by hospital. It is therefore important to understand whether studies of healthcare market structures that rely solely on traditional measures of quality are likely to be biased. Results (available on request) from a model without OHS change scores suggest that this is not the case: there is little change in the coefficients on the other quality variables. This is

Table 2

Estimated marginal utilities.

Variable	Est	SE
Main effects		
Change in Oxford Hip Score	0.130	0.008***
28-day emergency readmission rate (%)	−0.053	0.004***
1-year revision rate (%)	−0.017	0.010
28-day mortality rate (%)	−0.038	0.028
Distance (in km)	−0.270	0.007***
Distance ²	0.002	0.000***
Distance ³	−0.00001	0.000***
NHS trust – medium	−0.572	0.031***
NHS trust – multi-service	−0.663	0.096***
NHS trust – small	−0.832	0.039***
NHS trust – specialist	1.016	0.075***
NHS trust – teaching	−0.444	0.034***
Independent sector treatment centre	−1.564	0.039***
Primary care trust	−1.256	0.206***
Waiting time (proportion waiting >120 days)	−0.157	0.077*
Interaction with distance		
× Patient age	−0.002	0.000***
× Male	0.015	0.012
× Past utilisation	−0.018	0.005***
× Comorbidity count	0.0002	0.001
× Income deprivation	−0.261	0.046***
Interaction with waiting time		
× Patient age	0.008	0.003**
× Male	−0.076	0.055
× Past utilisation	0.022	0.063
× Comorbidity count	−0.046	0.033
× Income deprivation	1.126	0.469*
Interaction with change in Oxford Hip Score		
× Patient age	0.001	0.000*
× Male	−0.007	0.006
× Past utilisation	−0.007	0.007
× Comorbidity count	−0.008	0.003*
× Income deprivation	−0.420	0.050***
Interaction with 28-day emergency readmission rate		
× Patient age	−0.0004	0.000*
× Male	−0.001	0.004
× Past utilisation	0.012	0.004**
× Comorbidity count	0.001	0.002
× Income deprivation	0.122	0.027***
Interaction with 1-year revision rate		
× Patient age	0.001	0.000
× Male	0.012	0.009
× Past utilisation	−0.002	0.011
× Comorbidity count	0.000	0.005
× Income deprivation	−0.045	0.068
Interaction with 28-day mortality rate		
× Patient age	−0.001	0.001
× Male	−0.058	0.023*
× Past utilisation	0.046	0.028
× Comorbidity count	−0.008	0.016
× Income deprivation	−0.111	0.179
WTT (OHS change)	0.899	0.056***
WTT (Readmission rate)	−0.625	0.050***
WTT (Revision rate)	−0.067	0.041
WTT (Mortality rate)	−0.068	0.049
Number of patients	170,916	
Number of providers	228	
BIC	442,722	
Pseudo R ²	0.659	

Notes: Conditional logit model of choice of hospital for elective hip replacement patients treated between April 2010 and March 2013. Waiting time and quality metrics are lagged by one year. Coefficients are marginal utilities. Interactions of patient characteristics with distance², distance³ and provider type are not reported (available on request). WTT (OHS) is the ratio of the coefficient on the quality variable to the marginal utility of distance ($\beta_d + 2\beta_{d^2} + 3\beta_{d^3}$) evaluated at the average distance to the chosen provider.

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

²³ During our study period ISTCs were funded through block contracts and paid to provide care to a pre-specified number of NHS patients. However, most ISTCs did not fulfil their quotas although they generally had low waiting times (Naylor and Gregory, 2009). Our results are consistent with this observation and suggest a positive preference for public providers by NHS-funded hip replacement patients. Brown et al. 2015 also found evidence of a pro-public preference in the stated preferences of the general public in New Zealand.

²⁴ The average waiting time at provider level is 2.5 months, compared with around 8 months at the beginning of the century (Siciliani et al., 2014).

consistent with the low observed correlation between PROM change scores and the other quality metrics and suggests that the former pick up an additional different dimension of hospital quality that is valued by patients.

4.2.2. Patient heterogeneity

The coefficients on the interaction terms in Table 2 suggest that preferences vary across types of patient. We find, like other studies (Beckert et al., 2012; Propper et al., 2007), that older patients dislike distance more. They care less about waiting time and get greater marginal utility from improvements in the OHS change score and reductions in emergency readmissions. There is little difference between the preferences of male and female patients except that male patients have a greater dislike for providers with higher mortality. Preferences vary little by morbidity as measured by past emergency admissions or comorbidity count, except that more morbid patients have a greater dislike of distance and care less about OHS change rates. Finally, patients from neighbourhoods with greater income deprivation care more about distance and less about quality, although the estimated marginal utility of PROM quality is positive for all but the 1644 most deprived patients ($\leq 1\%$ of sample).

We use the detailed patient reported pre-operative OHS measures in our dataset to explore in more detail how patients in worse health status differ from their healthier counterparts in their sensitive to quality and willingness to travel for higher quality, as commonly assumed in the literature on hospital quality (Geweke et al., 2003; Gowrisankaran and Town, 1999). Sicker patients may have greater marginal disutility from distance but greater marginal utility from quality, so that they may be more or less willing to travel to a provider than health patients. The correlations between patients' pre-operative OHS and the routinely available morbidity measures are low, suggesting that they measure different aspects of the patient's condition at the time of admission.

The first model in Table 3 is the same as our baseline specification but with additional patient pre-operative OHS interactions. Interaction terms with other patient characteristics are omitted and available from the authors. We find that healthier patients are more willing to travel. Although the marginal utility from higher quality is unaffected by pre-operative OHS, healthier patients have a smaller disutility from distance and so are more willing to travel for higher quality.²⁵ Healthier patients are also more likely to choose a private provider, which is consistent with observed differences in patient mix across provider types (Browne et al., 2008).

The fact that pre-operative OHS data are available for only about 60% of patients raises concerns about response bias if unobserved factors affect propensity to respond and utility from providers.²⁶ To investigate if responders to the pre-operative PROM questionnaire have different preferences to non-responders we re-estimate the baseline specification of Table 2 for our full sample (responders and non-responders) but interact a dummy variable for responder status with all the main and interacted explanatory variables; pre-operative health status is not modelled. However, because the pre-operative PROM questionnaire is administered *after* the patient has chosen the provider, it

is unclear whether an observed response indicates differences in patient preferences or whether the choice determines the likelihood to observe a response. For example, private providers have higher response rates than NHS hospitals (Gomes et al., 2015; Gutacker et al., 2015b) and also tend to have higher observed quality and shorter waiting times. We address this concern by including the observed provider pre-operative response rate as a provider characteristic when modelling the choices of responders and non-responders. This variable is informative about the individual's propensity to fill in a pre-operative PROM questionnaire given the chosen provider²⁷.

In model 2 in Table 2 we add the provider response rate to the responder only model. Its coefficient is positive, which is to be expected since patients responding to the pre-operative questionnaire are more likely to have been treated at a provider with a higher response rate. Two effects are worth noting. The inclusion of the response rate reduces WTT for an increase in the OHS change score by about 8% relative to model 1, suggesting that providers' ability to administer the pre-operative questionnaire may be correlated with desirable, but unobserved, provider characteristics. More importantly, the estimated WTTs with or without adjusting (Table 2) for pre-operative health differ in a magnitude of only 100 metres. This suggests that the estimates from our baseline specification without adjustment for pre-operative health are robust to potential response bias.

Model 3 in Table 2 shows that responders and non-responders have generally similar revealed preferences, with the exception of preferences for waiting times (non-responders prefer lower proportions of long-waits), distance (non-responders are less willing to travel), revision rates (non-responders dislike providers with higher rates, whereas responders do not) and PROM quality (non-responders derive less utility from health gains and are thus less willing to travel for it). There is no difference with respect to the disutility from readmission rates or mortality.

4.3. Testing for unobserved provider characteristics

We also explore the possible impact of omitted hospital characteristics on our estimates of marginal utility for quality and other hospital characteristics. We first compare preferences of elective and emergency patients estimated from pooled choice models with a full set of emergency patient dummy variables interacted with all explanatory variables (see Section 3.2 for the detailed rationale). There are 73,496 emergency patients in our sample. Only 19.5% of emergency patients bypassed the nearest provider (see Appendix Fig. A1). Descriptive statistics for this patient group are reported in Appendix Table A4. Emergency patients' choice sets are the 50 closest providers who carried out hip replacement surgery on at least 30 emergency patients in the current and previous year. This rules out private and specialised providers who only treat elective hip replacement patients. 694 (0.9%) emergency patients were dropped because they attended a provider not in their choice set. All main effects are for patients with mean or modal elective patient characteristics.

We report results for two different specifications. The first model in Table 4 compares emergency patients with elective patients who choose NHS or independent providers (same as Table 2). However, there are some marked differences in observed characteristics between those two groups. For example, emergency patients are on average 12 years older than elective patients and have over twice as many recorded comorbidities. Hence in the second model reported in Table 5 we compare a set of elective and emergency

²⁵ In additional analyses we explored whether patients' past experience of care quality as measured by their individual change in OHS predicts their future willingness to return to the same provider for planned hip replacement of their other hip joint. We use data on 3195 patients that were treated in 2010/11, had pre- and post-operative OHS recorded and received a second primary hip replacement before the end of our dataset (31 March 2013). Of these, 2923 (91.4%) attended their initial provider for further surgery. Estimates from logit regression models show no association between the probability to re-attend and the experienced change in OHS from the initial surgery.

²⁶ We are not concerned about the implications of response rates for the hospital level case-mix adjusted OHS change scores as these have been shown to be robust to variations in response rate (Gomes et al., 2015).

²⁷ As a check, we first re-estimate the responder only model with the addition of provider pre-operative response rates but without interactions with pre-operative health. The results are robust to this sensitivity analysis, with the WTT of 0.9 km ($SE = 0.056$) for a standard deviation increase in PROM quality being similar to our baseline specification (full results available on request).

Table 3

Choice models allowing for patient pre-operative Oxford Hip Score.

Variable	Patients with pre-op OHS (1)		Patients with pre-op OHS (2)		All patients (3)					
					Responders (3a)		Non-responders (3b)		Difference (3c)	
	Est	SE	Est	SE	Est	SE	Est	SE	Est	SE
Main effects										
Change in Oxford Hip Score	0.153	0.009***	0.141	0.009***	0.139	0.009***	0.113	0.009***	0.026	0.009**
28-day emergency readmission rate (%)	−0.058	0.005***	−0.051	0.005***	−0.051	0.005***	−0.048	0.006***	−0.003	0.005
1-year revision rate (%)	−0.028	0.011*	−0.002	0.011	−0.003	0.011	−0.036	0.014**	0.032	0.013*
28-day mortality rate (%)	−0.137	0.033***	−0.052	0.030	−0.052	0.030	−0.021	0.034	−0.031	0.033
Distance (in km)	−0.265	0.008***	−0.264	0.008***	−0.265	0.009***	−0.281	0.006***	0.017	0.008*
Distance ²	0.002	0.000***	0.002	0.000***	0.002	0.000***	0.003	0.000***	0.000	0.000
Distance ³	−0.00001	0.000***	0.000	0.000***	0.000	0.000***	0.000	0.000***	0.000	0.000
NHS trust – medium	−0.517	0.035***	−0.591	0.034***	−0.583	0.034***	−0.565	0.036***	−0.018	0.031
NHS trust – multi-service	−0.948	0.125***	−0.961	0.124***	−0.951	0.123***	−0.625	0.107***	−0.325	0.126**
NHS trust – small	−0.799	0.043***	−0.854	0.043***	−0.849	0.042***	−0.885	0.046***	0.036	0.039
NHS trust – specialist	1.184	0.074***	0.961	0.076***	0.954	0.076***	0.979	0.100***	−0.025	0.068
NHS trust – teaching	−0.383	0.037***	−0.425	0.037***	−0.414	0.037***	−0.539	0.039***	0.125	0.034***
Independent sector treatment centre	−1.520	0.043***	−1.548	0.042***	−1.556	0.042***	−1.698	0.046***	0.141	0.042***
Primary care trust	−1.031	0.219***	−1.404	0.221***	−1.400	0.221***	−1.263	0.254***	−0.137	0.251
Waiting time (proportion waiting >120 days)	−0.343	0.088***	−0.026	0.086	−0.024	0.085	−0.399	0.092***	0.375	0.088***
Response rate (%)			0.019	0.001***	0.019	0.001***	−0.027	0.001***	0.046	0.001***
Interaction with pre-operative Oxford Hip Score										
× Change in Oxford Hip Score	0.000	0.000	0.000	0.000						
× 28-day emergency readmission rate (%)	0.000	0.000	0.000	0.000						
× 1-year revision rate (%)	0.000	0.001	0.001	0.001						
× 28-day mortality rate (%)	0.002	0.002	0.003	0.002						
× Distance (in km)	0.001	0.000***	0.001	0.000***						
× Distance ²	0.000	0.000**	0.000	0.000**						
× Distance ³	0.000	0.000**	0.000	0.000**						
× NHS trust – medium	−0.002	0.002	−0.004	0.002*						
× NHS trust – multi-service	−0.007	0.007	−0.008	0.007						
× NHS trust – small	−0.002	0.002	−0.003	0.002						
× NHS trust – specialist	0.014	0.003***	0.011	0.003***						
× NHS trust – teaching	−0.007	0.002***	−0.008	0.002***						
× Independent sector treatment centre	0.029	0.002***	0.027	0.002***						
× Primary care trust	0.000	0.008	−0.004	0.008						
× Waiting time (proportion waiting >120 days)	−0.001	0.005	0.002	0.005						
× Response rate (%)			0.000	0.000***						
WTT (OHS change)	1.075	0.062***	0.991	0.064***	0.974	0.063***	0.756	0.064***	0.218	0.062***
WTT (Readmission rate)	−0.689	0.057***	−0.612	0.056***	−0.604	0.056***	−0.544	0.063***	−0.060	0.060
WTT (Revision rate)	−0.116	0.046*	−0.006	0.045	−0.013	0.045	−0.139	0.053**	0.126	0.051*
WTT (Mortality rate)	−0.243	0.060***	−0.092	0.054	−0.093	0.054	−0.036	0.058	−0.057	0.057
Number of patients	113,846		113,846		170,916					
Number of providers	228		228		228					
BIC	296,526		292,634		432,899					
Pseudo R ²	0.658		0.662		0.667					

Notes: Conditional logit model of choice of hospital for elective hip replacement patients treated between April 2010 and March 2012. Waiting time and quality metrics are lagged by one year. Coefficients are marginal utilities. WTT is the ratio of the coefficient on the quality variable to the marginal utility of distance evaluated at the average distance to the chosen provider. Models in (1) and (2) are for patients reporting a pre-operation OHS. Model in (3) is for all patients and interacts a dummy variable for reporting a pre-operation OHS. Interaction effects are reported in (3c). All models also contain a full set of interactions of age, gender, past utilisation, Elixhauser comorbidities, and deprivation with hospital characteristics and distance (not reported). Standard errors are clustered at GP practice level.

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

patients matched exactly on age, gender, past emergency admissions, number of comorbidities, income deprivation and year of treatment. Additionally, we restrict the elective patient sample to those who used an NHS provider that treated at least 30 elective and emergency patients in that year, so that the choice sets are identical for elective and emergency conditional on location.

Both models suggest that OHS change scores have less influence on the provider used by emergency patients than for elective patient. The provider used by emergency patients is also less affected by more traditional readmission and mortality rates, though differences are statistically insignificant and less precisely estimated. In the second specification in Table 5, with closely matched patients, the estimated marginal utility of OHS changes ($\beta_q^{\text{emer}} = 0.058$) is just over one half of that for elective patients ($\beta_q^{\text{elec}} = 0.110$) and significant at $p < 0.05$. If we assume that emergency patients' demand is entirely inelastic to variation in observed

elective quality and that the estimated association for emergency patients is a result of omitted variables that affect emergency and elective patients in the same way, then the difference in the marginal utility of OHS changes ($\beta_q^A = 0.052$) can be interpreted as a lower bound estimate of the true effect of OHS change score on elective patient utility. The WTT for a one SD increase in OHS change scores then is 0.3 km (SE = 0.103), which is substantially smaller than the 0.9 km estimated in our main specification (see Table 2).²⁸

²⁸ Emergency patients are, on average, less willing to travel than elective patients. This is reflected by a more negative coefficient on the cubic distance term in Tables 4 and 5 so that utility decreases more rapidly as distance increases. Furthermore, emergency patients tend to be older than elective patients and age is associated with increased disutility from travel. This emphasises the need to compare WTTs for a defined reference patient, not at sample averages.

Table 4

Comparison of marginal utilities for elective and emergency patients.

Variable	Elective patients		Emergency patients		Difference	
	Est	SE	Est	SE	Est	SE
Change in Oxford Hip Score	0.130	0.008***	0.071	0.015***	0.058	0.014***
28-day emergency readmission rate (%)	−0.053	0.004***	−0.043	0.008***	−0.010	0.008
1-year revision rate (%)	−0.017	0.010	−0.063	0.022**	0.046	0.021*
28-day mortality rate (%)	−0.038	0.028	−0.011	0.057	−0.027	0.059
Distance (in km)	−0.270	0.007***	−0.324	0.011***	0.054	0.013***
Distance ²	0.002	0.000***	0.003	0.000***	0.000	0.000
Distance ³	0.000	0.000***	0.000	0.000***	0.000	0.000
NHS trust – medium	−0.572	0.031***	−0.660	0.051***	0.088	0.045*
NHS trust – multi-service	−0.663	0.096***	−1.021	0.167***	0.358	0.150*
NHS trust – small	−0.832	0.039***	−0.872	0.053***	0.040	0.048
NHS trust – specialist	1.016	0.075***	n/a		n/a	
NHS trust – teaching	−0.444	0.034***	−0.641	0.050***	0.197	0.047***
Independent sector treatment centre	−1.564	0.039***	n/a		n/a	
Primary care trust	−1.256	0.206***	−1.353	0.275***	0.097	0.183
Waiting time (proportion waiting >120 days)	−0.157	0.077*	−0.054	0.114	−0.103	0.107
WTT (OHS change)	0.899	0.056***	0.398	0.085***	0.501	0.076***
WTT (Readmission rate)	−0.625	0.050***	−0.405	0.080***	−0.219	0.075**
WTT (Revision rate)	−0.067	0.041	−0.203	0.072**	0.137	0.068*
WTT (Mortality rate)	−0.068	0.049	−0.016	0.081	−0.052	0.086
Number of patients	170,916		72,802			
Number of providers	228		137			
BIC	542,130					
Pseudo R ²	0.709					

Notes: Conditional logit model of choice of hospital for elective and emergency hip replacement patients treated between April 2010 and March 2013. Waiting time and quality metrics are lagged by one year. Coefficients are marginal utilities for the 'reference patient'. Elective and emergency patients are not matched on observed characteristics but the 'reference patient' in both patient populations is defined according to the average characteristics of the elective patient sample. WTT is the ratio of the coefficient on the quality variable to the marginal utility of distance evaluated at the median distance (in km). Model is estimated with a full set of dummy variables interacted with hospital characteristics and other interaction terms. All models also contain a full set of interactions of age, gender, past utilisation, Elixhauser comorbidities, and deprivation with hospital characteristics and distance (not reported). Standard errors are clustered at GP practice level.

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

Our second sensitivity check for possible bias due to unobserved hospital characteristics is to estimate our baseline specification (1) for elective patients with the addition of hospital fixed effects. In Table 6 in the model estimated with choice sets including ISTCs,

PROM quality still has a statistically significant effect on demand, whereas emergency readmission rates no longer do. The WTT to travel for PROM quality is however nearly 90% lower than that calculated from the results in Table 2 (0.1 km vs 0.9 km). Also, mortality

Table 5

Comparison of marginal utilities for elective and emergency patients – matched sample.

Variable	Elective patients		Emergency patients		Difference	
	Est	SE	Est	SE	Est	SE
Change in Oxford Hip Score	0.110	0.017***	0.058	0.016***	0.052	0.018**
28-day emergency readmission rate (%)	−0.060	0.010***	−0.041	0.009***	−0.019	0.011
1-year revision rate (%)	−0.099	0.025***	−0.078	0.025**	−0.021	0.030
28-day mortality rate (%)	−0.035	0.067	0.006	0.066	−0.041	0.087
Distance (in km)	−0.379	0.010***	−0.324	0.015***	−0.055	0.017**
Distance ²	0.005	0.000***	0.003	0.000***	0.002	0.000***
Distance ³	0.000	0.000***	0.000	0.000***	0.000	0.000***
NHS trust – medium	−0.817	0.054***	−0.642	0.055***	−0.176	0.061**
NHS trust – multi-service	−0.784	0.174***	−1.027	0.189***	0.242	0.220
NHS trust – small	−0.958	0.063***	−0.834	0.059***	−0.123	0.071
NHS trust – teaching	−0.489	0.058***	−0.622	0.054***	0.133	0.066*
Primary care trust	−1.419	0.324***	−1.509	0.366***	0.090	0.353
Waiting time (proportion waiting >120 days)	−0.546	0.137***	−0.200	0.131	−0.347	0.162*
WTT (OHS change)	0.617	0.094***	0.324	0.091***	−0.292	0.103**
WTT (Readmission rate)	−0.573	0.091***	−0.389	0.087***	0.184	0.106
WTT (Revision rate)	−0.323	0.081***	−0.254	0.082**	0.069	0.099
WTT (Mortality rate)	−0.049	0.095	0.008	0.094	0.058	0.123
Number of patients	31,631		31,631			
Number of providers	137		137			
BIC	97,740					
Pseudo R ²	0.794					

Notes: Conditional logit model of choice of hospital for elective and emergency hip replacement patients treated between April 2010 and March 2013. Waiting time and quality metrics are lagged by one year. Coefficients are marginal utilities for the 'reference patient'. Elective and emergency patients are matched exactly on observed characteristics (age, gender, past emergency utilisation in last year (none, once, or more), income deprivation of neighbourhood, number of Elixhauser comorbid conditions, year of treatment) and the 'reference patient' in both patient populations is defined according to the average (prior to matching) characteristics of the elective patient sample. Choice sets include only providers that treat at least 30 elective and 30 emergency hip replacement patient in this period. WTT is the ratio of the coefficient on the quality variable to the marginal utility of distance evaluated at the median distance (in km). Model is estimated with a full set of dummy variables interacted with hospital characteristics and other interaction terms. All models also contain a full set of interactions of age, gender, past utilisation, Elixhauser comorbidities, and deprivation with hospital characteristics and distance (not reported). Standard errors are clustered at GP practice level.

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

Table 6
Choice model controlling for unobserved time-invariant hospital effects.

	All providers (1)		NHS providers only (2)	
	Est	SE	Est	SE
Change in Oxford Hip Score	0.017	0.006**	0.014	0.007
28-day emergency readmission rate (%)	0.005	0.003	0.002	0.004
1-year revision rate (%)	0.003	0.010	0.001	0.012
28-day mortality rate (%)	0.045	0.022*	0.028	0.028
Distance (in km)	−0.293	0.007***	−0.334	0.010***
Distance ²	0.003	0.000***	0.003	0.000***
Distance ³	−0.00001	0.000***	−0.00001	0.000***
Waiting time (proportion waiting >120 days)	0.013	0.061	0.045	0.066
WTT (OHS change)	0.108	0.041**	0.077	0.042
WTT (Readmission rate)	0.050	0.037	0.014	0.040
WTT (Revision rate)	0.013	0.036	0.005	0.039
WTT (Mortality rate)	0.072	0.035*	0.040	0.040
Number of patients	170,916		146,839	
Number of providers	228		144	
BIC	394,881		246,842	
Pseudo R ²	0.698		0.762	

Notes: Conditional logit model of choice of hospital for elective hip replacement patients treated between April 2010 and March 2013. Waiting time and quality metrics are lagged by one year. Coefficients are marginal utilities. WTT is the ratio of the coefficient on the quality variable to the marginal utility of distance evaluated at the average distance to the chosen provider. Model in (1) does not impose restrictions on the type of provider in patients' choice sets. Model in (2) is based on a restricted choice set of NHS providers, thereby excluding patients that selected ISTCs. All models include indicator variables for hospitals (not reported). All models also contain a full set of interactions of age, gender, past utilisation, Elixhauser comorbidities, and deprivation with hospital characteristics and distance (not reported). Standard errors are clustered at GP practice level.

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

rates are now statistically significantly positively associated with demand. This may be the result of market entry of low-mortality ISTCs towards the end of our study period. Results are broadly similar when patients' choice sets are restricted to NHS hospitals, although we now find statistically insignificant coefficient estimates on all variables but distance. This is likely to be due to the limited within-provider variation over time in quality metrics and the fixed effects absorbing part of the effect of time-invariant quality on choice.

4.4. Public reporting

The public release of hospital quality information was intended to improve patients' knowledge about hospital quality and, thus, alter demand patterns. But patients and their GPs may already have had expectations about hospital quality (e.g. from the experience of other patients in the area) so that the new data may simply measure what was already known to them (Dranove and Sfekeas, 2008). We investigate this by exploiting a feature of the data collection process.²⁹ Information on patients' OHS was collected from April 2009 onwards but hospital PROM scores were not reported in the public domain until early 2010 due to the lag in post-operative data collection. Hence, choices of patients treated in the financial year 2009/10 (i.e. prior to our sample) could not have been influenced by the public release of OHS change scores for 2009/10.

²⁹ Chernew et al. (2008) and Dafny and Dranove (2008) both adopt a different approach and estimate Bayesian learning models to gauge patients' prior knowledge of health insurance quality and subsequent changes when report cards were released.

We estimate a conditional logit model for patients in 2009/10 and 2010/11 in which we assume that patient utility i from hospital j in year t is:

$$U_{ijt} = D'_{ij}\beta_{d,i} + D'_{ij}\beta_{d^2,i} + D'_{ij}\beta_{d^3,i} + Q'_{jt-1}\beta_{q,i} + OHS'_{j,2009/10}\beta_{OHS,i} + I_{t=2010/11}OHS'_{j,2009/10}\lambda_{OHS,i} + W'_{jt-1}\beta_{w,i} + Z'_{jt}\beta_{z,i} + v_{ijt} \quad (10)$$

This is the same as our model in (1) except that we assume that patients in 2009/10 choose as if they knew the OHS change scores for 2009/10 and patients in 2010/11 choose knowing the OHS change scores for 2009/10. Q'_{jt-1} denotes the other lagged quality variables (i.e. emergency readmission rates, mortality rates, revision rates) that were available to both patient groups.

We find that $\beta_{OHS} = 0.057$ (SE = 0.028) and $\lambda_{OHS} = -0.009$ (SE = 0.008). This suggests that patients in 2009/10 acted as if they knew the as yet unpublished OHS quality scores and that their choice behaviour was the same as that of patients choosing in 2010/11 when the OHS quality scores were published. Hence, the hospitals' OHS change scores may capture an aspect of quality which was known and valued by patients, and the publication of PROMs in the public domain does not seem to have altered this relation in the year of its introduction or subsequent years.

4.5. Responsiveness to auxiliary quality information

An alternative way to explore whether patients have an understanding and therefore respond to hospital quality is to test whether hospital choice is influenced by other publicly available quality metrics which are unlikely to be consulted as part of the choice process. Hip and knee replacements are typically performed by the same surgeons using common production inputs (e.g. same operation rooms, nursing input, diagnostic services, etc.). Hence, knee replacement-specific quality metrics may be informative about the quality of orthopaedic care provided since the correlation coefficient between the OHS and OKS change scores is 0.434.

We have re-estimated our main specification and included the change in Oxford Knee Score (OKS) as an additional regressor. Our sample is reduced to 169,391 hip replacement patients (550 provider-year observations) since not all providers have OKS change scores recorded and/or treat knee replacement patients. We find a positive and statistically significant coefficient on lagged OKS change scores of 0.121 (SE = 0.008). The coefficient on lagged OHS is slightly reduced to 0.088 (SE = 0.008), presumably due to the correlation between both metrics and the coefficient on lagged OKS picking up some of what was previously picked up by the OHS coefficient. All other coefficients are similar.

The significant coefficient on knee replacement quality in explaining choice of hip replacement provider could be because both knee replacement quality and hip replacement quality measures convey information to patients and their GP advisors about the quality of care they are likely to experience in the provider's orthopaedic department. Or, if we believe it is implausible that patients access information about knee replacement quality when choosing a hip replacement provider, we can interpret the coefficient as revealing to the econometrician that there are aspects of quality observable by patients, over and above those revealed in published quality metrics, which are correlated with knee replacement quality. In either case, the finding provides additional support for believing that patient choice of provider is influenced by quality.

5. The economic effects of quality on demand

We use the results from choice models to illustrate the effect of quality differentiation on hospital demand. Columns four and five of Table 7 provide the marginal utilities and the willingness to travel

Table 7
Effect sizes of hospital quality measures.

Quality indicator	Observed		Marginal utility	Effect of SD improvement in quality			Elasticity of demand
	Mean	SD		WTT	Demand change	% Demand change	
Baseline specification							
Change in Oxford Hip Score	19.8	1.4	0.130	0.9	35.3	9.8	1.4
Emergency readmission rate (in %)	5.7	2.4	0.053	0.6	24.5	6.8	0.2
Revision rate (in %)	0.9	0.8	0.017	0.1	2.6	0.7	0.01
Mortality rate (in %)	0.2	0.4	0.038	0.1	2.7	0.7	0.004
Time-invariant hospital fixed effects							
Change in Oxford Hip Score	19.8	1.4	0.017	0.1	4.6	1.2	0.02
Emergency readmission rate (in %)	5.7	2.4	0.005	−0.1	−2.1	−0.6	−0.01
Revision rate (in %)	0.9	0.8	0.003	−0.01	−0.5	−0.1	−0.002
Mortality rate (in %)	0.2	0.4	0.045	−0.1	−3.1	−0.8	−0.004
Elective vs emergency comparison							
Change in Oxford Hip Score	19.8	1.4	0.058	0.5	15.9	4.4	0.6
Emergency readmission rate (in %)	5.7	2.4	−0.010	0.2	4.8	1.3	0.03
Revision rate (in %)	0.9	0.8	0.046	−0.1	−7.3	−2.0	−0.02
Mortality rate (in %)	0.2	0.4	−0.027	0.1	1.9	0.5	0.003

SD = Standard deviation; MU = Marginal utility; WTT = Willingness to travel.

Notes: WTT is the ratio of the coefficient on the quality variable to the marginal utility of distance evaluated at the average distance to the chosen provider. Changes in volume and elasticities are averaged across hospital-year observations and are weighted by predicted demand $\hat{Y}_{ijt} = \sum_{ieMtu} P_{ijt}$. The WTT, effects of a SD improvement and elasticities for the three failure measures are expressed so that they reflect decreases in rates.

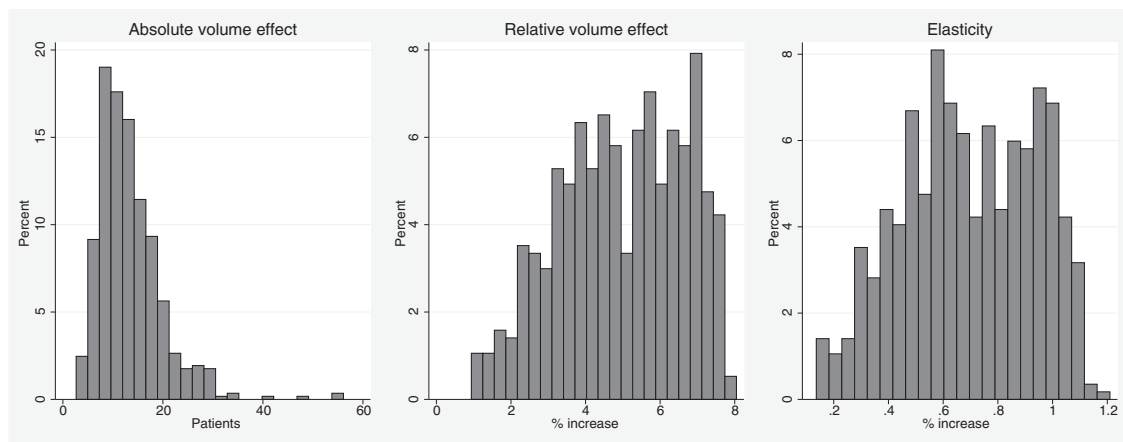


Fig. 2. Distribution of changes in hospital demand as a result of a SD increase in Oxford Hip Score change scores and quality elasticity of demand.

for one SD improvement in the quality measures. The sixth and seventh columns show the average total and relative change in demand from a one SD improvement in quality, and column eight gives the own quality demand elasticities. Quality improvements on the three failure measures correspond to decreases in rates. We provide calculations based on the baseline specification in Table 2, the hospital fixed effects specification in Table 6, and the specification comparing elective and emergency patients in Table 4. We focus our discussion on the latter since, under the assumption that elective and emergency patients have the same preferences over unobserved hospital characteristics, it accounts for unobservable hospital effects but is more precisely estimated than the fixed effects model. The marginal utilities used in these calculations are the estimated differences between elective and emergency patients.

The expected increase in demand for a one SD increase in OHS is 16 patients, or 4.4% of predicted demand at current quality levels. Decreases in mortality and emergency admission rates are associated with increases in demand, although the associations of these quality metrics with demand are not statistically significant. The marginal utility of quality as measured through revision rates is negative for both elective and emergency patients but more so for the latter so that quality improvements are associated with a negative effect

on demand (see Section 3.2). The effect of a one SD increase in OHS is larger than that of a one SD decrease in readmission rate.

There is substantial variation across providers in the effect of OHS change scores on own demand (Fig. 2). The number of additional patients arising from a one SD increase in OHS ranges from 3 to 56 (mean = 15.9). The relative volume increase from a one SD increase in OHS ranges from 1% to 8% (mean = 4.4%). The estimated elasticities range from 0.1 to 1.2 (mean = 0.6). About 42% of the variation in elasticities is explained by the amount of competition a provider faces, here measured by the Herfindahl–Hirschman Index (HHI).³⁰ Providers in more competitive areas (low HHI) face larger

³⁰ The HHI for provider j is calculated as the sum of the squared market shares of all providers $j^* = 1, \dots, J^*$ that service LSOA $a = 1, \dots, A$, here denoted as s_{aj} , weighted by the proportion of the provider's observed total activity originating from this LSOA, s_{ja} , so that:

$$HHI_j = \sum_a s_{ja} * \left[\sum_{j^*} (s_{aj^*})^2 \right] \quad (11)$$

Hospital catchment areas are defined as all LSOAs within a radius of 30 km around the hospital.

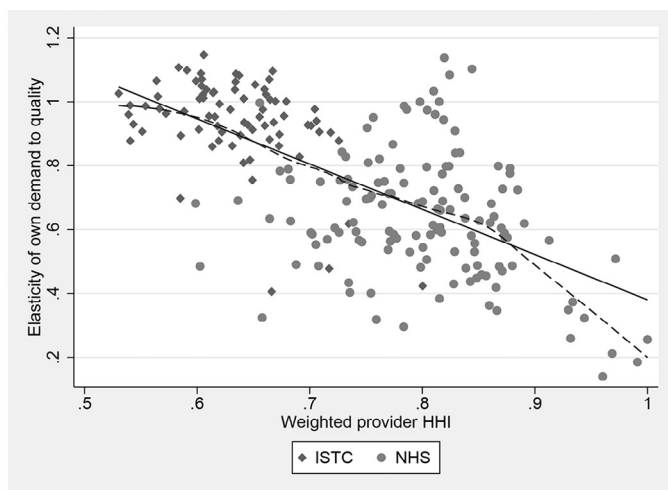


Fig. 3. Differences in quality elasticity of demand between providers in competitive (low HHI) and non-competitive (high HHI) markets. Notes: Solid line shows best linear fit (Intercept = 1.80 (SE = 0.08), slope = -1.42 (SE = 0.10), $R^2 = 0.42$). Dashed line shows LOWESS curve.

quality elasticities than those in less competitive areas (high HHI), with elasticities falling by approximately 0.14 per 0.1 increase in HHI (assuming a linear effect; $p < 0.001$) (Fig. 3). Markets are more competitive in areas where independent sector treatment centres are active. However, this result should be interpreted cautiously since the functional form of the choice model does not allow patients' preferences for quality to vary by provider type.

We also examine the effect of changes in the quality of other providers on a provider's demand. Higher cross-quality demand elasticities make it more likely that increases in one provider's quality will trigger an increase in the quality of other providers. Fig. 4 shows how cross-quality elasticities decline rapidly as the distances between providers increase. Whereas a 1% increase in a competitor's PROM quality is associated with a -0.32% reduction in demand if the competitor is located within 10 km, this falls to -0.11% when the competitor is 30 km away. Hence, quality competition is likely to be restricted geographically.

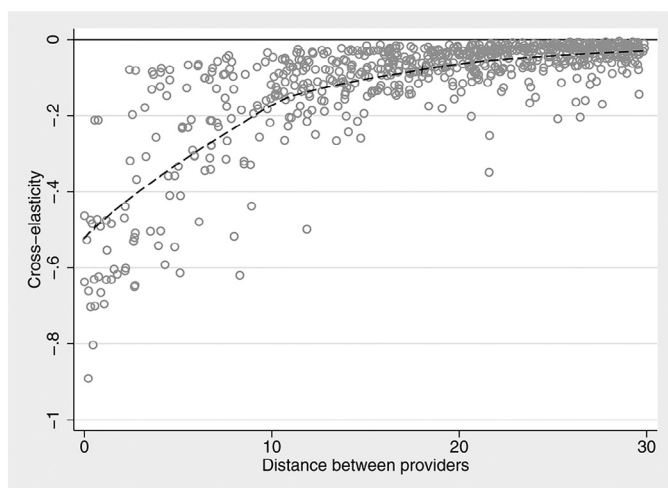


Fig. 4. Percentage change in demand as a result of percentage change in competitor's quality. Notes: Dashed line shows LOWESS curve.

6. Discussion and concluding remarks

This paper is the first to test whether patient choice of hospital is influenced by hospital quality as measured by the change in health status and functioning rather than by traditional failure measures such as readmission and mortality rates. We use data on observed choices for all NHS-funded hip replacement patients treated between April 2010 and March 2013 in public and private hospitals in England.

The average health gain, as measured by the PROMs-based OHS change score, is only weakly correlated with the simple condition-specific mortality, revision and emergency readmission rates across hospitals. This suggests that choice models which do not include PROMs-based measures may miss important factors affecting patient choice of hospital. This is borne out by the results from our choice models. We find that elective hospital demand responds to observed quality as measured by PROMs after conditioning on patient characteristics, distance to hospital and other quality metrics. Hospital demand is relatively more responsive to OHS change quality than emergency readmission rates. There is no statistically significant association of choice of provider with mortality, though this may be because the mortality risk for an elective hip replacement is very small, nor was there an association with revision rates after primary hip replacement surgery. Finally, we found that patients may respond to hospital quality even when corresponding information is not made publicly available, yet.

Our results are relevant for policy. First, they show that providers can attract additional patients by raising quality, which is a necessary condition for competition to improve quality. Second, we find that demand is more responsive to quality measures based on the change in patient health status due to treatment, rather than to crude measures such as rates of mortality or emergency readmission. Thus hospitals wishing to attract patients can do so by improving aspects of quality with a more immediate link to outcomes experienced by all patients. Third, the public release of information on hospital quality may not in itself trigger a change in demand if patients have other channels by which they can infer hospital quality. Fourth, our results emphasise the importance of market structure in determining incentives for competing on quality: the effect of quality changes on the providers' ability to attract patients away from rival hospitals decreases rapidly with distance. Thus some providers with few nearby rivals face little incentive to raise quality. This is likely to be true of all types of competition, since the number of rivals is likely to be in part determined by the number of potential patients, but is not always acknowledged in the health policy debate about stimulating quality improvement efforts through increased competition. Of course, whether or not providers engage in quality competition based on published quality information depends on whether they *perceive* their demand to be elastic to quality changes and on how much they value their reputation. We cannot answer these questions with our data. Finally, our estimated effect may, at least in part, be driven by unobserved hospital characteristics correlated with observable quality. We attempted to allow for this. Using emergency hip replacement patients as a control group who should not be affected by elective quality suggests that the willingness to travel for higher PROM quality could be up to 50% smaller. A model including hospital fixed effects suggested that willingness to travel could as much as 90% smaller. These alternative models either rely on strong assumptions or are poorly identified because quality scores did not vary much over time. Nevertheless, they suggest that the 10% increase in demand for a one standard deviation increase in PROM quality implied by our baseline model is likely to be an upper bound estimate.

We also explore whether patient preferences vary according to observed and unobserved patient characteristics. We find that the preference for PROM quality increases with age and decreases with income deprivation and comorbidity burden. Qualitatively similar

results are obtained for preferences for quality as approximated by emergency readmission rates. We find that patient reported pre-operative Oxford Hip Score is only weakly correlated at patient level with other measures of pre-operative morbidity, such as recorded morbidities or number of previous admissions, which can be extracted from routine hospital records and are commonly used in the literature. We do not find evidence that preferences for quality vary with pre-operative health status as reported by the patient herself. But because healthier patients derive less disutility from travel, they have, *ceteris paribus*, a higher willingness to travel for quality. This finding may help shed light on the underlying mechanism of the 'distance bias' that has been described in other studies of hospital demand (Geweke et al., 2003; Gowrisankaran and Town, 1999).

There remains scope for further research. First, we cannot tell whether the estimated effect of quality is driven by patients' choices or choices by their GPs or by bargaining between patients and their GPs. Who is taking the decision has implications for the design of policies intended to influence choice of provider. Second, we did not investigate whether patients undergoing knee replacement, groin hernia repair or varicose vein surgery, for which PROMs-based quality measures are also available, are responsive to quality. These patients differ from hip replacement patients. For example, knee replacement patients are generally younger and, possibly, more willing and able to search for hospital quality information online. Future work could explore difference in demand patterns across treatments. Third, our findings may be specific to the condition under study. Patients undergoing surgery with considerable risk of mortality may be more sensitive to quality information since the health costs of choosing an inferior provider would likely be more significant. For example Gaynor et al. (2012) study choices made by coronary bypass surgery patients and find that they are sensitive to procedure-specific hospital mortality rates. Hence, the choice context may affect both the general sensitivity to hospital quality as well as the importance of individual failure measures such as emergency readmission or death; both of which are rare after hip replacement. Fourth, there may be scope for quasi-experimental studies using hospitals that undertook quality improvement plans and comparing them with hospitals that did not. Any changes in quality scores observed in the former group but not the latter are likely to reflect genuine improvements, not statistical noise, and would allow a more precise estimation of the effect of quality on demand. Fifth, the role of other quality measures, such as reputation or overall patient experience and satisfaction, could be explored when they become available.³¹

In conclusion, our results suggest that patients' choice of hospital is influenced by the health gain from treatment, not just by the likelihood of extreme, rare events such as death or an emergency readmission.

Acknowledgements

We are grateful for comments and suggestions from Alistair McGuire, Alastair Gray, Rowena Jacobs, Carsten Crede, Andrew Street, Karen Bloor, Richard Cookson, the editor and two anonymous referees as well as those received during presentations at the Universities of Birmingham, Newcastle, York and Melbourne, the King's Fund, the 2nd Annual BECCLE Competition Policy Conference, the 5th Network of Industrial Economists Doctoral Colloquium, and the Leeds 2015 meeting of the Health Economic Study Group. This research was funded by the Department of Health in England under the Policy Research Unit in the Economics of Health and Social Care Systems (Ref 103/0001). The views expressed are those of the authors and not necessarily those of the funders.

³¹ In 2013 the NHS launched a "Friends and Family Test" asking patients whether they would recommend their hospital to friends or family.

Hospital Episode Statistics are copyright © 2009–2015, re-used with the permission of The Health & Social Care Information Centre. All rights reserved.

Appendix

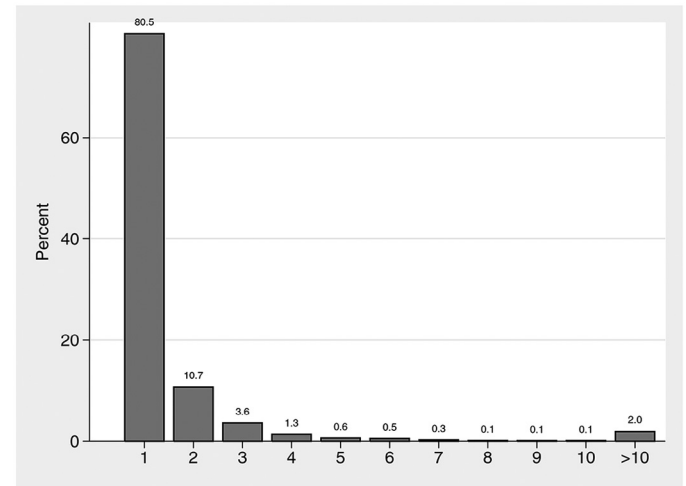


Fig. A1. Percentage of emergency patients who went to their Nth nearest hospital.

Table A1
Mixed logit choice model.

Variable	Mean		Standard deviation	
	Est	SE	Est	SE
Change in Oxford Hip Score	0.130	0.008***	−0.001	0.001
28-day emergency readmission rate (%)	−0.053	0.004***	−0.001	0.001
1-year revision rate (%)	−0.017	0.010	−0.001	0.001
28-day mortality rate (%)	−0.039	0.028	−0.019	0.069
Distance (in km)	−0.270	0.007***		
Distance ²	0.002	0.000***		
Distance ³	0.000	0.000***		
NHS trust – medium	−0.572	0.031***		
NHS trust – multi-service	−0.663	0.096***		
NHS trust – small	−0.832	0.039***		
NHS trust – specialist	1.016	0.075***		
NHS trust – teaching	−0.444	0.034***		
Independent sector treatment centre	−1.564	0.039***		
Primary care trust	−1.256	0.206***		
Waiting time (proportion waiting >120 days)	−0.157	0.077*		
WTT (OHS change)	0.899	0.056***		
WTT (Readmission rate)	−0.625	0.050***		
WTT (Revision rate)	−0.067	0.041		
WTT (Mortality rate)	−0.068	0.048		
Number of patients	170,916			
Number of providers	228			
BIC	442,785			

Notes: Random coefficient (mixed) multinomial logit model of choice of hospital for elective hip replacement patients treated between April 2010 and March 2013. Waiting time and quality metrics are lagged by one year. Coefficients are marginal utilities. Random coefficients are specified for OHS change, readmission rate, revision rate and mortality rate. WTT is the ratio of the coefficient on the quality variable to the marginal utility of distance evaluated at the average distance to the chosen hospital (in km). Interaction terms with patient characteristics are not reported (available on request). Standard errors are clustered at GP practice level. The mean coefficients do differ from those reported in Table 2 for the conditional logit model if there is less rounding.

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

Table A2
Sensitivity analyses.

Variable	Model (1)		Model (2)		Model (3)		Model (4)	
	Est	SE	Est	SE	Est	SE	Est	SE
Change in PROM score	0.104	0.008***	0.132	0.008***	5.253	0.301***	0.138	0.011***
28-day emergency readmission rate (%)	−0.052	0.005***	−0.054	0.004***	−0.057	0.004***	−0.055	0.006***
1-year revision rate (%)	−0.026	0.012*	−0.016	0.010	−0.019	0.011	−0.065	0.015***
28-day mortality rate (%)	−0.098	0.033**	−0.038	0.028	−0.032	0.028	0.039	0.036
Distance (in km)	−0.265	0.006***	−0.270	0.007***	−0.270	0.007***	−0.315	0.011***
Distance ²	0.002	0.000***	0.002	0.000***	0.002	0.000***	0.003	0.000***
Distance ³	0.000	0.000***	0.000	0.000***	0.000	0.000***	0.000	0.000***
NHS trust – medium	−0.569	0.034***	−0.573	0.031***	−0.584	0.032***	−0.715	0.038***
NHS trust – multi-service	−0.705	0.094***	−0.653	0.095***	−0.671	0.094***	−0.696	0.099***
NHS trust – small	−0.850	0.040***	−0.833	0.039***	−0.836	0.040***	−0.954	0.044***
NHS trust – specialist	1.021	0.074***	1.043	0.075***	1.064	0.076***	1.083	0.086***
NHS trust – teaching	−0.517	0.035***	−0.457	0.034***	−0.406	0.034***	−0.510	0.039***
Independent sector treatment centre	−1.736	0.062***	−1.487	0.048***	−1.486	0.038***		
Primary care trust	−1.017	0.198***	−1.273	0.206***	−1.212	0.205***	−1.349	0.219***
Waiting time	−0.376	0.091***	0.028	0.015	−0.220	0.077**	−0.312	0.090***
WTT (PROM change)	0.738	0.059***	0.915	0.056***	0.861	0.051***	0.832	0.065***
WTT (Readmission rate)	−0.622	0.058***	−0.629	0.050***	−0.663	0.050***	−0.564	0.062***
WTT (Revision rate)	−0.106	0.050*	−0.063	0.041	−0.075	0.044	−0.228	0.053***
WTT (Mortality rate)	−0.176	0.060**	−0.068	0.048	−0.057	0.049	0.060	0.056
Number of patients	120,377		170,916		169,621		146,839	
Number of providers	230		228		223		144	
BIC	331,496		442,697		432,611		282,596	
Pseudo R ²	0.637		0.659		0.661		0.726	

Notes: Conditional logit model of choice of hospital for elective hip replacement patients treated between April 2010 and March 2013. Waiting time and quality metrics are lagged by one year if not otherwise stated. Coefficients are marginal utilities. WTT is the ratio of the coefficient on the quality variable to the marginal utility of distance evaluated at the average distance to the chosen hospital. Interaction terms with patient characteristics are not reported (available on request). Standard errors are clustered at GP practice level. Model (1) – Waiting time and quality metrics are contemporaneous. Based on observed choices for patients treated between April 2009 and March 2012. Since April 2012 PROM scores have been reported separately for primary and revision hip replacement surgeries so that our measures of PROM quality are no longer comparable. Model (2) – Median waiting time substituted for proportion of patients waiting longer than 120 days (both lagged). Model (3) – Lagged EQ-5D change scores substituted for lagged OHS change scores. Model (4) – Patients' choice sets exclude independent sector treatment centres.

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

Table A3
Estimated willingness to travel in different choice sets.

	30closest		50closest		100closest	
	Est	SE	Est	SE	Est	SE
WTT (PROM change)	0.907	0.055***	0.899	0.056***	0.904	0.057***
WTT (Readmission rate)	−0.604	0.050***	−0.625	0.050***	−0.642	0.051***
WTT (Revision rate)	−0.062	0.040	−0.067	0.041	−0.072	0.042
WTT (Mortality rate)	−0.069	0.048	−0.068	0.049	−0.070	0.049

Notes: Conditional logit model of choice of hospital for elective hip replacement patients treated between April 2010 and March 2013. Waiting time and quality metrics are lagged by one year. Willingness to travel (WTT) is the ratio of the coefficient on the quality variable to the marginal utility of distance evaluated at the average distance to the chosen hospital. Standard errors are clustered at GP practice level. Full results available on request.

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

Table A4
Descriptive statistics – emergency sample.

Variable	Obs	Mean	SD	ICC
Patient characteristics				
Distance travelled (in km)	73,496	14.2	27.1	
Distance travelled past closest provider (in km)	73,496	4.2	25.3	
Number of providers within 10 km radius	73,496	1.0	1.2	
Number of providers within 30 km radius	73,496	5.0	4.7	
Age	73,496	80.9	9.8	
Male	73,496	0.3	0.4	
Past utilisation	73,496	0.7	1.2	
Number of Elixhauser conditions	73,496	1.0	1.6	
Income deprivation	73,496	0.1	0.1	
Provider characteristics				
Observed volume	393	187.2	87.0	80.7%
Waiting time (proportion waiting > 120 days)	393	0.2	0.1	43.7%
Change in Oxford Hip Score	393	19.4	1.3	47.9%
28-day emergency readmission rate (%)	393	5.97	2.22	37.0%
1-year revision rate (%)	393	0.95	0.72	22.4%
28-day mortality rate (%)	393	0.20	0.25	4.5%

Obs = Observations; SD = Standard deviation; ICC = Intraclass correlation coefficient. Notes: Patient characteristics for emergency patients treated between April 2010 and March 2013. Waiting time and quality metrics are for financial years 2009/10 to 2011/12. Provider characteristics are unweighted. The intraclass correlation coefficient (ICC) measures the proportion of variation that occurs between providers, rather than over time.

References

- Appleby, J., Devlin, N., 2004. Measuring success in the NHS: using patient assessed health outcomes to manage the performance of health care providers. The King's Fund. London.
- Arthroplasty Clinical Outcomes Registry, 2015. 2014 Annual report. <http://www.acornregistry.org/images/ACORN_AnnualReport_2014.pdf> (accessed 06.11.15).
- Basch, E., Torda, P., Adams, K., 2013. Standards for patient-reported outcome-based performance measures. *JAMA: The Journal of the American Medical Association* 310 (2), 139–140.
- Beckert, W., Christensen, M., Collyer, K., 2012. Choice of NHS-funded hospital services in England. *The Economic Journal* 122, 400–417.
- Besley, T., Ghatak, M., 2003. Incentives, choice, and accountability in the provision of public services. *Oxford Review of Economic Policy* 19 (2), 235–249.
- Brekke, K., Gravelle, H., Siciliani, L., Straume, O., 2014. Patient choice, mobility and competition among health care providers. In: Levaggi, R., Montefiori, M. (Eds.), *Health Care Provision and Patient Mobility*. Springer.
- Brooks, R., 1996. EuroQol: the current state of play. *Health Policy* 37, 53–72.
- Brown, P., Panattoni, L., Cameron, L., Knox, S., Ashton, T., Tenbensel, T., Windsor, J., 2015. Hospital sector choice and support for public hospital care in New Zealand: results from a labeled discrete choice survey. *Journal of Health Economics* 43, 118–127. doi:10.1016/j.jhealeco.2015.06.004. forthcoming.
- Browne, J., Jamieson, L., Lewsey, J., van der Meulen, J., Copley, L., Black, N., 2008. Case-mix & patients' reports of outcome in independent sector treatment centres: comparison with NHS providers. *BMC Health Services Research* 8 (1), 78.

- Bundorf, M., Chun, N., Goda, G., Kessler, D., 2009. Do markets respond to quality information? The case of fertility clinics. *Journal of Health Economics* 28 (3), 718–727.
- Burns, L., Wholey, D., 1992. The impact of physician characteristics in conditional choice models for hospital care. *Journal of Health Economics* 11, 43–62.
- Chernew, M., Gowrisankaran, G., Scanlon, D.P., 2008. Learning and the value of information: evidence from health plan report cards. *Journal of Econometrics* 144, 156–174.
- Chou, S.-Y., Deily, M.E., Li, S., Lu, Y., 2014. Competition and the impact of online hospital report cards. *Journal of Health Economics* 34, 42–58.
- Cutler, D., Huckman, R., Landrum, M., 2004. The role of information in medical markets: an analysis of publicly reported outcomes in cardiac surgery. *The American Economic Review* 94, 342–346.
- Dafny, L., Dranove, D., 2008. Do report cards tell consumers anything they don't already know? The case of Medicare HMOs. *The RAND Journal of Economics* 39 (3), 790–821.
- Dawson, J., Fitzpatrick, R., Carr, A., Murray, D., 1996. Questionnaire on the perceptions of patients about total hip replacement. *The Journal of Bone and Joint Surgery. British volume* 78-B, 185–190.
- Department of Health, 2008. Guidance on the routine collection of Patient Reported Outcome Measures (PROMs). The Stationary Office, London.
- Department of Health, 2012. Patient Reported Outcome Measures (PROMs) in England: The case-mix adjustment methodology. The Stationary Office, London.
- Dranove, D., Sfeka, A., 2008. Start spreading the news: a structural estimate of the effects of New York hospital report cards. *Journal of Health Economics* 27 (5), 1201–1207.
- Dranove, D., Kessler, D., McLellan, M., Satterthwaite, M., 2003. Is more information better? The effects of 'report cards' on health care providers. *The Journal of Political Economy* 111 (3), 555–588.
- Elixhauser, A., Steiner, C., Harris, D., Coffey, R., 1998. Comorbidity measures for use with administrative data. *Medical Care* 36 (1), 8–27.
- Epstein, A., 2010. Effects of report cards on referral patterns to cardiac surgeons. *Journal of Health Economics* 29 (5), 718–731.
- Etzioni, D., Fowl, R., Wasif, N., Donohue, J., Cima, R., 2013. Distance bias and surgical outcomes. *Medical Care* 51 (3), 238–244.
- Faber, M., Bosch, M., Wollersheim, H., Leatherman, S., Grol, R., 2009. Public reporting in health care: how do consumers use quality-of-care information? A systematic review. *Medical Care* 47 (1), 1–8.
- Folland, S., 1983. Predicting hospital market shares. *Inquiry: A Journal of Medical Care Organization, Provision and Financing* 20 (34–44).
- Gaynor, M., Seider, H., Vogt, W., 2005. The volume-outcome effect, scale economies, and learning by doing. *The American Economic Review* 95 (2), 243–247.
- Gaynor, M., Propper, C., Seiler, S., 2012. Free to choose? Reform and demand response in the English National Health Service. Working Paper 18574. National Bureau of Economic Research.
- Geweke, J., Gowrisankaran, G., Town, R.J., 2003. Bayesian inference for hospital quality in a selection model. *Econometrica: Journal of the Econometric Society* 71 (4), 1215–1238.
- Gomes, M., Gutacker, N., Bojke, C., Street, A., 2015. Addressing missing data in patient-reported outcome measures (PROMs): implications for comparing provider performance. *Health Economics* 25 (5), 515–528. doi:10.1002/hec.3173. forthcoming.
- Gowrisankaran, G., Town, R.J., 1999. Estimating the quality of care in hospitals using instrumental variables. *Journal of Health Economics* 18 (6), 747–767.
- Gravelle, H., Dusheiko, M., Sutton, M., 2002. The demand for elective surgery in a public system: time and money prices in the UK National Health Service. *Journal of Health Economics* 21 (3), 423–449.
- Gravelle, H., Santos, R., Siciliani, L., 2014. Does a hospital's quality depend on the quality of other hospitals? A spatial econometrics approach. *Regional Science and Urban Economics* 49, 203–216.
- Gutacker, N., Bojke, C., Daidone, S., Devlin, N., Street, A., 2013. Hospital variation in patient-reported outcomes at the level of EQ-5D dimensions: evidence from England. *Medical Decision Making: An International Journal of the Society for Medical Decision Making* 33 (6), 804–818.
- Gutacker, N., Bloor, K., Cookson, R., 2015a. Comparing the performance of the Charlson/Deyo and Elixhauser comorbidity measures across five European countries and three conditions. *European Journal of Public Health* 25 (Suppl. 1), 15–20.
- Gutacker, N., Street, A., Gomes, M., Bojke, C., 2015b. Should English healthcare providers be penalised for failing to collect patient-reported outcome measures (PROMs)? *Journal of the Royal Society of Medicine* 108 (8), 304–316.
- Health & Social Care Information Centre, n.d. Estates Returns Information Collection (ERIC) 2014/15. <<http://hefs.hscic.gov.uk/Downloads%5CHEFS%20Data%20Definitions%202014-15.pdf>> (accessed 08.05.16).
- Health & Social Care Information Centre, 2013a. Emergency readmissions to hospital within 28 days of discharge: primary hip replacement surgery. *Compendium of Population Health Indicators*. <https://indicators.ic.nhs.uk/download/NCHOD/Specification/Spec_33D_533ISP4CPP1_12_V1.pdf> (accessed 24.11.14).
- Health & Social Care Information Centre, 2013b. Finalised Patient Reported Outcome Measures (PROMs) in England: April 2011 to March 2012. <<http://www.hscic.gov.uk/catalogue/PUB11359/final-proms-eng-apr11-mar12-fin-report-v2.pdf>> (accessed 26.03.15).
- Hensher, D.A., Greene, W.H., 2003. The mixed logit model: the state of practice. *Transportation* 30, 133–176.
- Ho, K., 2006. The welfare effects of restricted hospital choice in the US medical care market. *Journal of Applied Econometrics* 21 (7), 1039–1079.
- Hodgkin, D., 1996. Specialized service offerings and patients' choice of hospital: the case of cardiac catheterization. *Journal of Health Economics* 15 (3), 305–332.
- Hole, A.R., 2007a. A comparison of approaches to estimating confidence intervals for willingness to pay measures. *Health Economics* 16 (8), 827–840.
- Hole, A.R., 2007b. Fitting mixed logit models by using maximum simulated likelihood. *The Stata Journal* 7 (3), 388–401.
- Howard, D., 2005. Quality and consumer choice in health care: evidence from kidney transplantation. *Topics in Economic Analysis and Policy* 5 (1), 1349.
- Hunt, L.P., Ben-Shlomo, Y., Clark, E.M., Dieppe, P., Judge, A., MacGregor, A.J., Tobias, J.H., Vernon, K., Blom, A.W., 2013. 90-day mortality after 409,096 total hip replacements for osteoarthritis, from the National Joint Registry for England and Wales: a retrospective analysis. *The Lancet* 382 (9898), 1097–1104.
- Judge, A., Chard, J., Learmonth, I., Dieppe, P., 2006. The effects of surgical volumes and training centre status on outcomes following total joint replacement: analysis of the Hospital Episode Statistics for England. *Journal of Public Health* 28 (2), 116–124.
- Jung, K., Feldman, R., Scanlon, D., 2011. Where would you go for your next hospitalization? *Journal of Health Economics* 30 (4), 832–841.
- Landelijke Registratie Orthopedische Implantaten, 2013. LROI Report 2013: Insight into Quality & Safety. <<http://www.lroi.nl/uploads/KL/6J/KL6JSE01JL2yLbJHtNatQ/LROI-rapportage-2013-Engels.pdf>> (accessed 06.11.15).
- Luft, H.S., Hunt, S.S., Maerki, S.C., 1987. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? *Health Services Research* 22 (2), 157–182.
- Luft, H.S., Garnick, D., Mark, D., Peltzman, D., Phibbs, C., Lichtenberg, E., McPhee, S., 1990. Does quality improve choice of hospital. *JAMA: The Journal of the American Medical Association* 263, 2899–2906.
- Marshall, M.N., Romano, P.S., Davies, H.T.O., 2004. How do we maximize the impact of the public reporting of quality of care? *International Journal for Quality in Health Care: Journal of the International Society for Quality in Health Care/ISQua* 16 (S1), 57–63.
- Martin, S., Smith, P.C., 1999. Rationing by waiting lists: an empirical investigation. *Journal of Public Economics* 71 (1), 141–164.
- Martin, S., Smith, P.C., 2003. Using panel methods to model waiting times for National Health Service surgery. *Journal of the Royal Statistical Society. Series A, (Statistics in Society)* 166 (3), 369–387.
- McConnell, K., Lindrooth, R., Wholey, D., Maddox, T., Bloom, N., 2016. Modern management practices and hospital admissions. *Health Economics* 25 (4), 470–485.
- McFadden, D., 1974. Conditional logit analysis of qualitative choice behaviour. In: Zarembka, P. (Ed.), *Frontiers in Economics*, vol. 4. Academic Press, New York, pp. 105–142.
- McGrail, K., Bryan, S., Davis, J., 2012. Let's all go to the PROM: the case for routine patient-reported outcome measurement in Canadian healthcare. *HealthcarePapers* 11, 8–18.
- McLennan, D., Barnes, H., Noble, M., Davis, J., Garrett, E., Dibben, C., 2011. The English Indices of Deprivation 2010. Department for Communities and Local Government.
- Moya, L., Piatti, A., Pecoraro, V., Ricci, C., Virgili, G., Salanti, G., Germagnoli, L., Liberati, A., Banfi, G., 2012. Timing matters in hip fracture surgery: patients operated within 48 hours have better outcomes. a meta-analysis and meta-regression of over 190,000 patients. *PLoS ONE* 7 (10), e46175.
- Monstad, K., Engester, L., Espehaug, B., 2006. Patient preferences for choice of hospital. Tech. rep. *Health Economics Bergen Working Paper No 05/06*.
- Moscone, F., Tosetti, E., Vittadini, G., 2012. Social interaction in patients' hospital choice: evidence from Italy. *Journal of the Royal Statistical Society. Series A, (Statistics in Society)* 175 (2), 453–472.
- National Patient Safety Agency, n.d. NRLS Data: What are the NRLS 'cluster'? <www.nrls.npsa.nhs.uk/EasySiteWeb/getresource.axd?AssetID=135267> (accessed 08.05.16).
- Naylor, C., Gregory, S., 2009. Independent sector treatment centres. The King's Fund Briefing.
- Neuburger, J., Hutchings, A., van der Meulen, J., Black, N., 2013. Using patient-reported outcomes (PROs) to compare the provider of surgery: does the choice of measure matter? *Medical Care* 51 (6), 517–523.
- NICE, 2011. Hip Fracture: The Management of Hip Fracture in Adults, vol. 124. London. NICE Guidelines.
- Pope, D., 2009. Reacting to rankings: evidence from "America's best hospitals". *Journal of Health Economics* 28 (6), 1154–1165.
- Propper, C., Damiani, M., Leckie, G., Dixon, J., 2007. Impact of patients' socioeconomic status on the distance travelled for hospital admission in the English National Health Service. *Journal of Health Services Research & Policy* 12 (3), 153–159.
- Rolfson, O., Kärrholm, J., Dahlberg, L., Garellick, G., 2011. Patient-reported outcomes in the Swedish Hip Arthroplasty Register. *The Journal of Bone and Joint Surgery. British volume* 82-B (7), 867–875.
- Ruwaard, S., Douven, R., 2014. Quality and hospital choice for cataract treatments: the winner takes most. CPB Netherlands Bureau for Economic Policy Analysis.
- Santos, R., Gravelle, H., Propper, C., 2015. Does quality affect patients' choice of doctor? Evidence from England. *The Economic Journal* doi:10.1111/econj.12282. forthcoming.
- Siciliani, L., Moran, V., Borowitz, M., 2014. Measuring and comparing health care waiting times in OECD countries. *Health Policy* 118 (3), 292–303.
- Sivey, P., 2008. The effect of hospital quality on choice of hospital for elective heart operations in England (Ph.D. thesis). University of York.
- Sivey, P., 2012. The effect of waiting time and distance on hospital choice for English cataract patients. *Health Economics* 21, 444–456.

- Tay, A., 2003. Assessing competition in hospital care markets: the importance of accounting for quality differentiation. *The Rand Journal of Economics* 34 (4), 786–814.
- Train, K.E., 2003. *Discrete Choice Methods with Simulation*, 1st ed. Cambridge University Press, Cambridge.
- Varagunam, M., Hutchings, A., Black, N., 2015. Relationship between patient-reported outcomes of elective surgery and hospital and consultant volume. *Medical Care* 53 (4), 310–316.
- Varkevisser, M., Geest, S.A., Schut, F.T., 2010. Assessing hospital competition when prices don't matter to patients: the use of time-elasticities. *International Journal of Health Care Finance and Economics* 10 (1), 43–60.
- Varkevisser, M., van der Geest, S.A., Schut, F.T., 2012. Do patients choose hospitals with high quality ratings? Empirical evidence from the market for angioplasty in the Netherlands. *Journal of Health Economics* 31 (2), 371–378.
- Vrangbaek, K., Robertson, R., Winblad, U., Van de Bovenkamp, H., Dixon, A., 2012. Choice policies in Northern European health systems. *Health Economics, Policy and Law* 7 (S1), 47–71.
- Wang, J., Hockenberry, J., Chou, S.-Y., Yang, M., 2011. Do bad report cards have consequences? Impact of publicly reported provider quality information on the CABG market in Pennsylvania. *Journal of Health Economics* 30 (2), 392–407.
- Wooldridge, J.M., 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, Massachusetts.