Early work in progress: Please do not cite or circulate without permission

Identifying differences in performance when choice of provider is endogenous: application of the Hausman-Taylor estimator to unbalanced panel data

Alex Turner, Silviya Nikolova, Matt Sutton

Manchester Centre for Health Economics, Institute of Population Health, The University of Manchester

Address for correspondence: alexander.turner@manchester.ac.uk

Abstract

Background: Patients in England can receive treatment for NHS elective procedures either at privately-operated independent sector treatment centres (ISTCs) or in the public NHS. In line with the intention of ISTC contracts, patients' assignment to providers is not exogenous. ISTCs differ from NHS providers in terms of patient case-mix, workforce profile, and the regulatory framework they operate in. None of the previous studies comparing ISTCs and NHS providers has allowed for this endogeneity in patient assignment.

Study Question: To compare patient reported outcomes after elective surgery for hip replacement in independent sector treatment sector centres (ISTCs) and NHS providers allowing for endogeneity.

Methods and Data: While we partially control for the difference between providers by introducing a patient-invariant indicator for provider type, a part of this difference likely remains unexplained. Accounting for the unbalanced panel nature of the data, we use a range of methods (Random Effects, Mundlak and Hausman-Taylor) to control for the correlation between the unobserved provider effect and the explanatory variables. We apply these to a sample of 32,463 hip replacement patients admitted for treatment in the financial years 2009/10 and 2010/11, and recorded in the Patient Reported Outcome Measures (PROMs) data set.

Results: Patients treated at ISTCs were healthier and had less severe symptoms prior to treatment. After risk-adjusting and accounting for provider endogeneity we find larger improvements in outcomes for ISTC patients. However, depending on the assumptions made regarding the exogeneity of different types of covariates, the magnitude and significance of the differences change. The ISTC effect can increase up to a factor of 2.2 compared to the Random Effects case, which assumes exogeneity of all covariates with respect to the unobserved provider effect.

Conclusions: There is a lack of evaluation of the effects of ISTCs on patient outcomes. As the NHS moves towards greater involvement of the independent sector in the provision of health services, it is important to develop an appropriate methodological framework to ensure that their performance is correctly evaluated.

[328 words]

Funding: This work was funded by the MRC Methodology Research Programme Grant, G0901491 "Attribution of patient reported outcomes to the effects of care providers".

1 Introduction

Private sector involvement in the English National Health Service (NHS) has been gradually increasing over the past two decades, moving from a role as suppliers of construction services, medicines, and information technologies, to providers of primary and secondary care. Creation of the internal market, introduced across England in 1991, opened the door for private sector involvement in secondary care. This involved a split of purchasers from providers within the NHS, with the idea that both price and quality competition between providers would drive efficiency improvements and improve the quality of care. However, this also allowed NHS purchasers to buy care from non-NHS providers. Although the incoming Labour government of 1997 rejected the internal market, recognition of a role for the private sector remained. The NHS plan, published in 2000, emphasised cooperation between the public and private sector, announcing a national framework for partnership between the two (Department of Health, 2000). Details of this were formally introduced in a concordat between the Department of Health and the Independent Health Care Association, the trade association for independent sector healthcare providers. This aimed to formalise the association between the two sectors in order to reduce the *ad hoc* spot-purchasing of treatments from the private sector, in favour of a more long-term relationship whereby the private sector could help the NHS meet the needs of patients (Department of Health and Independent Health Care Association, 2000). Furthermore, reforms to increase patient choice, which gave patients greater control over where elective surgery was conducted, including options for treatment at private or overseas hospitals at the expense of the NHS, reintroduced competition for admissions between providers in the two sectors (Department of Health, 2002a). This was introduced explicitly in 2008, and evidence indicates that a growing proportion of the revenue created by private sector providers are generated by NHS patients who came to them through free choice (Sussex, 2009).

The most notable result of these reforms has been the participation of Independent Sector Treatment Centres (ISTCs) in the provision of elective surgery and diagnostic procedures to NHS-funded patients. NHS procurement of services from ISTCs was announced in 2002, with an aim of increasing NHS capacity (Department of Health, 2002b), reducing waiting times for routine procedures, and encouraging innovation in services (King's Fund, 2009). The Department of Health offered contracts to ISTCs in two waves. Treatment of NHS patients in ISTCs began in October 2003, and by May 2008 36 ISTC contracts were in place; 26 from "Wave 1" and 10 from "Phase 2" (Sussex, 2009). In 2007/8 ISTC activity represented approximately 2% of total elective activity, but this figure is much higher in particular specialties such as orthopaedics and ophthalmology (King's Fund, 2009).

The growth of the private sector in health care provision is a contentious issue and, given that the rollout of ISTCs has contributed substantially to this in the UK, many concerns over its implementation have been raised. In particular, issues regarding differential fees paid to ISTCs and the differential costs they incur have generated much antipathy (Bardsley and Dixon, 2011). NHS providers are paid on a fee-for-service basis, receiving a fixed tariff per procedure. The contracts offered to ISTCs differed from this arrangement. In "Wave 1", ISTCs were offered five-year contracts which guaranteed payment for a set number of procedures,

irrespective of whether these were all carried out. In addition, the per-procedure tariff paid to ISTCs was approximately 11.2% above NHS equivalent assuming the full number of contracted procedures was met. In reality, only 85% of contracted procedures were carried out, implying that the tariff premium for ISTCs was closer to 30% (Sussex, 2009). ISTCs offered contracts in Wave 1 also did not have to train clinical staff and thus experienced a cost advantage over NHS providers. Likewise, it was stated that the casemix of ISTC patients would be significantly less costly than those admitted to NHS providers (Sussex, 2009)[1]. More recent research has since confirmed this (Browne et al., 2008; Chard and Kuczawski, 2011; Mason et al., 2010). This has led to concerns that ISTCs had been "cherry picking" less complex patients (Bardsley and Dixon, 2011). However it was always the aim that the NHS would treat more complex and severe patients, with more routine cases being treated at ISTCs due to their lack of intensive care facilities (Pérotin et al., 2013; Sussex, 2009). Contracts offered in "Phase 2" were similar to those in "Wave 1" but fees paid to ISTCs were only guaranteed for a portion of the total number of contracted procedures. If this portion of procedures were carried out, per-procedure tariffs would average out at 7.2% lower than NHS costs (Sussex, 2009). However, these premiums could be borne out of necessity. These additional payments may have been mandatory to cover entry costs to ensure market participation from ISTCs (Pérotin et al., 2013), and required to induce ISTCs to invest in the capacity desired by the Department of Health (Sussex, 2009). ISTCs also experience comparatively higher market exit costs (Sussex, 2009). Additionally, ISTCs are subject to the same national tariffs as NHS providers once their initial five-year contracts end, meaning concerns regarding this issue are likely to be short term.

However, the most important ISTC-related concern has been over the relative quality of care being provided in private hospitals. These concerns have been driven by a perception of lower quality of staff at independent sector sites. Labour force restrictions placed on ISTCs, which prevented them from recruiting NHS staff in the first six months of their five-year contracts, led to a large proportion of staff being recruited from overseas. Although doctors and surgeons operating in ISTCs must be registered with the General Medical Council, as is the case in NHS providers, they are not subject to the scrutiny of the Advisory Appointments Committees which serve as an additional quality control mechanism for workers in the NHS (King's Fund, 2009)[2]. Many believe that this over-reliance on foreign healthcare workers could lead to poorer quality of care (Bardsley and Dixon, 2011). A recent study has aimed to clarify the theoretical predictions of whether it is private or public provision which generates the greatest level of care quality, and adapts these predictions for the case of ISTC versus NHS quality (Pérotin et al., 2013). They base their argument on the model developed in Hart et al. (1997), which argues that if contracts offered to for-profit organisations for the provision of government services are fee-for-service in nature and have no stipulation about a minimum level of quality, as is the case with ISTC contracts, then the incentive to cost minimise will outweigh those to improve quality, resulting in an under-supply of quality. However, it also argues that if reducing costs does not impact quality, then the level of quality supplied by the for-profit sector will outstrip that supplied by the not-for-profit sector. Despite this,

[1] It must also be noted that ISTCs did experience cost disadvantages compared to NHS providers, with the latter benefiting from state aid in many areas such as pension costs (Sussex, 2009).

[2] A recent investigation found that, although there were shortfalls in quality in some areas, ISTC recruitment processes met minimum requirements (Healthcare Commision, 2007).

they also include a caveat that if public sector employees derive significant benefits from higher levels of quality, akin to the intrinsic motivation hypothesis, then public provision will provide higher levels of quality. Given that studies on whether public sector employees are intrinsically more motivated is mixed, then no concrete theoretical predictions can be made as to whether private or public providers will supply higher levels of quality (Pérotin et al., 2013).

Empirical evidence on this topic is vast and has produced mixed results. Pérotin et al. (2013) conclude that findings from studies comparing clinical quality across for-profit, non-profit, and state-owned hospitals have proved inconclusive. However, studies into the difference in care quality between ISTCs and NHS providers are relatively scant. Early literature suggests that this has been mainly due to the poor nature of data collection in ISTCs, meaning that comparison of the outcomes between patients treated at these sites with those treated in NHS providers has not been viable (Health Committee, 2006; Healthcare Commision, 2007; King's Fund, 2009; Mason et al., 2010; National Centre for Health Outcomes Development, 2005). Of the recent literature, one study found that quality of life improvements and improvements in functional status were significantly higher in ISTCs compared to NHS providers for hip & knee replacement and cataract surgery, but lower for patients undergoing hernia repair. However, the authors cited weak casemix-adjustment and the low number of ISTCs participating as reasons not to place too much weight on their findings (Browne et al., 2008). A later study comparing outcomes for hip & knee replacement, varicose veins surgery, and hernia repair conducted a similar study but with improved levels of ISTC participation. Using patient reported outcomes data from 2008 and 2009, it compared outcomes between 72 NHS providers and 25 ISTCs. Contrary to predictions, it found that, for measures of quality of life, functional status and complication rates, patients treated at ISTCs had statistically significantly better outcomes than those treated at NHS providers, although this was only found for hip & knee replacement, with no difference in outcomes found for hernia repair and varicose vein surgery (Chard and Kuczawski, 2011). However, not all of the NHS providers and ISTCs who were invited to participate took part in the study. Additionally, patients treated at both NHS and independent sector providers were not required to participate in the study. If the providers which participated differed systematically from providers which did not, the patients who chose to participate differ systematically from those who did not, and these systematic differences are correlated with provider type, then these results may be subject to considerable selection bias. Furthermore, the authors also highlight that they fail to study patient satisfaction. Given that a recent review found that outcomes reported by patients could be correlated with satisfaction (Healthcare Commision, 2007), not accounting for this could have also biased results. A recent review also states that results from this study cannot be used to assess whether ownership (public or private) affects treatment-related outcomes, citing that it does not take into account other provider characteristics that could affect outcomes, such as volume of admissions and the degree of provider specialism (Bardsley and Dixon, 2011).

These criticisms highlight considerable methodological flaws in many of the previous attempts to assess whether provider type (ISTC or NHS provider) affects the quality of care. In particular the vast majority fail, possibly due to limitations of the data, to casemix-adjust adequately. Furthermore, ordinary least squares (OLS) is employed in the majority of cases, meaning that many characteristics of providers unobserved to the

researcher cannot be controlled for. They also fail to account for selection into provider type due to both Department of Health decisions that ISTCs should treat less complex patients, and patient selection as a result of the initiatives to increase patient choice.

The most recent study on this topic (Pérotin et al., 2013) stands as a lone exception. They estimate a switching regression which accounts for both patient selection into provider type and unmeasured provider-level heterogeneity, to examine whether outcomes for patients treated at ISTCs differ significantly from those treated at NHS providers using aggregated non-clinical indices of quality. These indices included "access", which covers uncertainty over admission date, waiting time prior to admission, and waiting time for a bed once admitted; and "comfort", which measures noise levels, cleanliness, quality of food, and whether healthcare workers treated them with respect and dignity. The results offer mixed evidence on the direction of an ownership effect, with ISTCs performing better in some domains of quality, and NHS providers performing better on others. They aggregate these 'domain-specific' effects into a single overall effect, and find that ownership does not have a statistically significant effect on the quality of care.

In this paper, we study data collected from hip replacement patients covered by the national patient reported outcomes (PROMs) programme, and apply a range of methods designed for use in panel data to overcome the methodological shortfalls in the previous literature. The national PROMs initiative has, since 2008, required all providers treating NHS patients to record data on a range of health and non-health variables for patients undergoing four key elective admissions. Recording data for all patients in all providers in this manner mitigates the problems with selection encountered in Chard and Kuczawski (2011). The data set-up here is theoretically different from cases where panel data methods are usually applied. Instead of observing multiple units, for example individuals, at different time periods, we instead observe multiple units (in this case providers), repeated over another set of units (in this case patients). Additionally, instead of separating the error term into a unit error (i.e. an individual error) and a time-varying random error, we separate the error into a fixed provider error and a patient-varying random error.

We apply traditional random effects and fixed effects models, along with models introduced in Mundlak (1978) and Hausman and Taylor (1981). These models differ in the assumptions made about the correlation between explanatory variables and the error terms, the efficiency of estimates, and whether the effects of fixed provider characteristics are identified.

2 Data

2.1 PROMs-HES linked data

We use patient-level data from the national patient reported outcome measures (PROMs) programme for the financial years 2009/10 and 2010/11. Since 1st April 2009, the PROMs programme has collected data on a wide range of health and treatment-related outcomes for all patients in England undergoing four key elective

interventions paid for by the NHS[3]: hip replacement; knee replacement; inguinal hernia repair; and varicose veins surgery. We focus only on hip replacement in this study. The national PROMs programme collects data through two questionnaires. The first is a pre-treatment survey, which is administered to patients either on the day of admission in person by hospital staff or at any time in the interval between a patient being passed fit for surgery and the intervention taking place (Department of Health, 2008). Post-treatment data are collected by the NHS Information Centre via a postal survey approximately six months after surgery (London School of Hygiene and Tropical Medicine, 2007). National PROMs data are linked at the patient level to Hospital Episode Statistics (HES) data. Data on socio-economic characteristics are obtained from the Neighbourhood Statistics website and linked at lower super output area (LSOA) level[4].

We restrict the sample to include only those patients where data are available for the dependent variable and all explanatory variables (described below). The final linked dataset contains 32,463 patient episodes.

2.2 Dependent variable

The PROMs questionnaire provides data on a series of health metrics on which we can compare NHS providers and ISTCs. Here we focus on the Oxford Hip Score (OHS), collected at the post-operative questionnaire, as our sole barometer of provider performance. The OHS measures symptom severity for hip problems by assessing functional problems and pain through a series of 12 questions (Dawson and Fitzpatrick, 1996). Responses to these questions are recorded on a four-point scale and are summed to create an overall score ranging from 0 (worst health state) to 48 (best health state).

2.3 Explanatory variables

Our main variable of interest, ISTC status, is constructed using provider type data in HES. Patients were coded as treated at an ISTC if their provider type was identified as either an independent sector treatment centre or a treatment centre at an independent sector site. The richness of the HES and PROMs datasets allows us to employ an extensive risk-adjustment procedure. The PROMs survey collects data on a wide range of health metrics at both the pre-operative and post-operative questionnaires, and the majority of the pre-operative health measures are used as covariates. This includes a categorical measure of general health, which rates a patient's self-perceived general health on a five-point scale ranging from 1(excellent) to 5(poor); a dummy for whether a patient considers themselves as disabled, a range of dummies on the presence of a range of self-

[3] This excludes patients who pay privately for their hip or knee replacements. These accounted for approximately one-third of hip and replacements in England and Wales in 2008 (Dr Foster Health, 2009).
[4] These are areas which cover an average population of approximately 1500 individuals.

reported comorbidities[5]; the EQ-VAS, which rates a patient's self-perceived health on a 100-point visual scale ranging from 0 (worst possible health) to 100 (best possible health); and categorical variables for each of the dimensions used in the construction of the EQ-5D index[6]. Using HES data we also control for the general severity of a patient using the Charlson index (Charlson and Pompei, 1987). Comorbidities are placed into 17 "Charlson categories" based on their ICD-10 codes. Weights are then assigned to each category, with weights increasing in the severity of the comorbidities in each category. The sum of these weights across all conditions for a particular patient provides the severity index for that patient.

We also control for a patient's age, included up to the fifth polynomial and standardised to avoid problems with collinearity, gender, ethnic group, living arrangements, length of symptoms prior to treatment, whether they have had a previous surgery on the same hip, whether they are undergoing a revision surgery, and their level of income deprivation. We use the income domain of the English Index of Multiple Deprivation 2010 (IMD), which assigns an income deprivation score to a patient depending on the proportion of individuals in their lower super output area (LSOA) who reside in households which are reliant on one or more means-tested benefits (Department for Communities and Local Government, 2011). Whether a patient is assisted in the completion of the post-operative questionnaire is also controlled for. Most importantly, we control for the value of the OHS recorded in the pre-operative questionnaire. Given that the dependent variable is the post-operative OHS, including this means that we can interpret the coefficient corresponding to the ISTC variable as the average difference in health improvement from treatment from being treated at an ISTC relative to being treated at an NHS provider.

A set of provider characteristics, other than ISTC status, are also controlled for, as recommended in Bardsley and Dixon (2011). We control for the size of provider using two measures of volume. Using HES data we record the total volume of patients treated by a provider over the course of our two-year sample period as a measure of the general size of a provider. We also record the total number of patients treated at each provider for the conditions covered by the PROMs dataset. This is used as a measure of the size of elective surgery within each provider. As a measure of scope we also record the number of HRGs covered by the treatments conducted at each provider. This can also be viewed as an indicator of provider specialisation.

[5] These include heart disease, high blood pressure, stroke, circulation problems, lung disease, diabetes, kidney disease, diseases of the nervous system, liver disease, depression, and arthritis.

[6] Patients are asked to rate problems in 5 dimension of health (mobility, self care, usual activities, pain/discomfort, and anxiety/depression) on a 3 point scale ("no problems", "some or moderate problems", "extreme problems").

3 Methods

3.1 Model set-up

In the PROMs dataset we have J providers. Provider j is observed over $\mathrm{i} = 1, \ldots, \mathrm{N_j}$ patients. Given that providers do not treat identical numbers of patients this implies that the PROMs data set has an unbalanced panel profile. The model specification can be expressed as follows:

$$y_{i,j} = \beta X_{i,j} + \gamma Z_j + \alpha_j + \epsilon_{i,j}$$

Where $y_{i,j}$ is the health outcome of patient i following hip replacement surgery at provider j, which could be a public, NHS hospital or private, independent sector treatment centre (ISTC). β and γ are vectors of coefficients associated with patient-varying $X_{i,j}$ and patient-invariant provider characteristics Z_j. The patient-level random component $\epsilon_{i,j}$ is assumed uncorrelated with (X, Z, α). The unobserved provider effect α_j is a provider-specific random variable, which is distributed independently across providers, with variance σ_α^2.

Given the model specification above, there are three "classic" scenarios for which $E(\varepsilon_{i,j}|X, Z) = 0$, $V(\varepsilon|X, Z) = \sigma_\varepsilon^2 I$ and all consistency-related assumptions hold for X, Z, ε. Under the most restrictive scenario, $\alpha_j = 0$, and OLS should be used. In the second case, it is assumed that $\alpha_j \sim (0, \sigma_\alpha^2)$, $E(\alpha_j|X, Z) = 0, Cov(\epsilon_{i,j}, \alpha_k) = 0$. Under this set of assumptions, random effects should be employed. If any of the α-related assumptions in Case 2 are violated, there is a tendency to use fixed effects methods. The last method will eliminate any type of provider-specific endogeneity, heterogeneity, or correlation between α_j components, but comes at the expense of losing information. In addition, as our primary effect of interest is the (possibly) endogenous assignment of patients to public or private providers, the fixed effects estimation sweeps away the effect of ISTC and any provider specific characteristics, such as hospital staffing levels or number of beds, which may influence the choice of provider.

However, if we know exactly what Case 2 assumptions are the violated, we might end up with a more efficient estimator than the fixed effects (Case 3) estimator. Two popular methods that could be used for endogenous provider effects were introduced in Mundlak (1978) and in Hausman and Taylor (1981).

Mundlak (1978) assumes that all explanatory variables are related to individual effects and parameterises the endogenous provider effects using averages of all patient-level explanatory variables over the panel dimension. Hausman and Taylor (1981) method is applicable to situations when there is correlation between subsets of X and Z and the unobserved provider-specific effect α_j.

The Hausman-Taylor method allows us to identify such patient-varying and provider-level characteristics that are uncorrelated with the unobserved provider effect α_j and obtain consistent estimates of β and γ. This method is an IV approach with a smart selection of instruments. The identification of parameters comes from variables within the model. It uses the fact that mean-deviation variables are, by construction, exogenous with respect to provider-specific endogeneity. There are three types of instruments: all mean-deviations, means of

exogenous patient-varying X, exogenous provider-specific Z. All mean-deviations are used to identify β. Means of exogenous patient-varying X are used to identify γ and β. Finally, the exogenous provider-specific Z are used to identify γ and β as well. Assuming only provider-specific endogeneity exists, the method is more efficient than fixed effects as it uses more observations to estimate the model.

Given selection into provider type, the indicator for treatment in an ISTC is assumed to be endogenous. Also, given that it was Department of Health policy that less severe patients would be treated at ISTCs, then we also treat all baseline health indicators as endogenous. All other explanatory variables are assumed to be exogenous.

In brief, for the identification of endogenous provider-level covariates and efficient estimation of β and γ:

- The random effects estimator is the best if there is no correlation between unobservable provider effects α and covariates patient-varying X or provider-level Z or between patient-level error ϵ and X or Z.
- If the true functional relationship between the dependent variable Y and covariates includes patient-varying X and their averages, then Mundlak estimates are the best.
- If the true functional relationship between the dependent variable Y and covariates includes patient-varying X and their averages $\bar{X}$, and either the provider-level averages $\bar{X}$ or provider-level covariates Z are correlated with provider-level unobservables α, the provider-level averages of exogenous X along with exogenous Z and mean-deviations of X can be used as instruments and the Hausman-Taylor method must be employed.

Particular issues arise when applying the Hausman-Taylor model in the case of an unbalanced panel. Gardner (1998) and Allin (2011) state that adjustments must be made to the instrument set in such circumstances. In the traditional Hausman-Taylor model, unweighted provider averages are used as instruments. Gardner argues that this is correct in the balanced case when the GLS transformation, θ, is constant across providers. However in the unbalanced case, θ differs between providers, and thus θ-weighted provider averages must be used for instruments. Allin (2011) raises similar concerns to Gardner (1998) but proposes a different change to the instrument set, stating that the quasi-demeaned exogenous patient varying variables must also be included as instruments.

3.2 Specification Tests

We perform a series of specification tests. First we would like to know whether panel techniques are more appropriate than standard OLS. We test for the presence of fixed errors using the Breusch-Pagan Lagrange multiplier test. The null hypothesis is that the provider-level variance component of the error term is zero. Under the alternative, OLS-implied statistical inference is incorrect. A random effects model is preferable in

this case as long as provider-specific heterogeneity is not correlated with any of the explanatory variables. We apply tests developed in Hausman (1978) to examine this. Failure to accept the null implies that there exists some type of endogeneity in the data and we need instruments to estimate the model. We test the consistency of Hausman-Taylor estimates using a Hausman specification test which compares these estimates to those from a fixed effects model. This implicitly tests whether the assumptions made regarding which explanatory variables are exogenous and endogenous are correct. Failure to accept the null can be the result of incorrect specifications of which variables are endogenous, instrument endogeneity, weak instruments or a more general form of patient-level error endogeneity i.e. a failure of independence assumption regarding explanatory variables and the random patient-level error.

4 Results

4.1 Summary statistics

The sample statistics (Table 1) indicate that that a lower proportion of males are treated at ISTCs compared to NHS providers (40.6% vs 41.8%). The same is true for the ethnicity variable, with ISTCs treating a lower proportion of non-white patients (0.7% vs 1.6%). ISTC patients are slightly older on average (68.1 vs 67.9), and are less income deprived. Furthermore, a lower proportion of ISTC patients report having being assisted with completion of the post-operative questionnaire (4.8% vs 7.7%). Living arrangements are generally consistent between ISTC patients and patients treated at NHS hospitals, although a slightly greater proportion of ISTC patients live with a partner, spouse, family member or friend (75.1% vs 73.4%) and a slightly lower proportion live alone (25.0% vs 26.0%). ISTCs also treat a lower proportion of patients having revision surgery (2.7% vs 8.2%), maybe due to their increased difficulty compared to primary procedures. Consistent with the aims of the Department of Health, the sample statistics indicate that ISTC treat less severe patients. This is consistent across all binary measures of health with a lower proportion of ISTC patients reporting disability (45.3% vs 60.5%), having had a previous surgery (4.5% vs 10.4%), and prevalence of all self-reported comorbidities. The categorical measure of general health indicates a similar picture, with a greater proportion of ISTC patients reporting excellent general health (7.2% vs 4.6%) and a smaller proportion reporting poor general health (1.2% vs 3.9%). The EQ-5D variables tell the same story, with ISTC patients reporting higher average health on the EQ-VAS (69.2 vs 65.0), and a greater (smaller) proportion of ISTC patients reporting having no (severe) problems on each of the EQ-5D health dimensions. The figures for the Charlson index agree, with ISTC patients having a much lower average severity score (0.2 vs 0.347). ISTC patients also experienced shorter length of symptoms prior to surgery, with this group having a lower proportion of patients with symptoms greater than 10 years (6.7% vs 8.9%) and a greater proportion with patients with symptoms less than 1 year (17.5% vs 13.7%).

4.2 Specification tests

Results of the specification tests are reported in Table 2. The Breusch-Pagan test for the existence of a provider effect in the error term rejects the null of no provider effect, with a p-value < 0.001. This implies that some type of error component model (random effects, fixed effects, or HT model) should be used, rather than OLS. The Hausman test, which tests whether random effects or fixed effects model are preferable, gives an unexpected result. The null of the consistency of the random effects estimates is not rejected (p-value=0.434), indicating a lack of correlation between any unobserved provider-level heterogeneity and the explanatory variables. Thus results of the specification tests suggest that a random effects approach should be employed. Nevertheless, we conduct the test of whether the HT model be used over the fixed effects model. This test fails to reject the null that HT estimates are consistent (p-value=1.000), and thus the HT model should be used over the fixed effects model, due to its greater efficiency and ability to identify the effects of all explanatory variables.

4.3 Regression results

Estimates from all of the models are presented in Table 3. Estimates from the preferred random effects specification indicate that being treated an ISTC significantly increases the health improvement from treatment compared to being treated at an NHS provider, with ISTC patients improving, on average, 1.651 points more on the OHS scale than NHS provider patients. This effect is identical in sign and significance to the OLS estimate, although is greater in magnitude. As provider type is fixed over patients treated within each provider, an ISTC effect is not identified in the fixed effects specification. Including provider averages of the patient-varying variables allows this effect to be identified in the Mundlak model. This effect is much lower in magnitude and becomes statistically insignificant. The magnitude of the effect of being treated in an ISTC is much greater in the HT specification compared the random effects estimate. However, due to this coefficient being so imprecisely estimated this effect is not statistically significant at any significance level.

Random effects estimates for the patient-varying covariates and estimates from all other specifications are identical in terms of sign and significance, and are remarkably similar in magnitude. Across the majority of health metrics, being in poorer health prior to treatment reduces the OHS reported post-operatively. Self-reported cancer and the mobility and usual activities dimensions of the EQ-5D are exceptions; with poorer pre-operative health on these metrics having a significant positive effect on patient's post-operative OHS. Results also indicate that patients reporting having had a previous surgery on the same hip, or recorded as having a revision surgery have significantly lower levels of improvement on the OHS. This indicates that repeat surgeries do not have as much potential to generate health improvements. OHS improvements also reduce significantly with age and the duration of symptoms prior to surgery. There also exists considerable inequalities in health improvement based on socio-economic status, ethnicity and gender, with both IMD income deprivation and

being non-white having a negative and statistically significant effect on improvements in the OHS, and being male having a positive and statistically significant effect. Additionally, patients reporting to have been assisted with the completion of the post-operative questionnaire also report significantly lower OHS improvements. Being assisted with questionnaire completion may be another indicator of health: those with lower levels of health are more likely to need assistance.

Like the ISTC effect, estimates for the patient-invariant provider characteristics vary between specifications. Estimates suggest that the number of HRGs covered by a provider has a positive effect, although this is insignificant in all models. Although total volume of admissions has no effect, the volume of PROMs patients treated at each hospital is found to have a statistically significant positive effect on OHS improvement in all models. As with the ISTC estimates, the HT estimates are larger in magnitude but estimated less precisely.

5 Conclusion

5.1 Discussion

This paper is the first to apply the Hausman-Taylor model to a new type of data set-up, where instead of individuals or providers being observed over time, providers are observed over multiple patients. In this set-up Hausman-Taylor models, which generate consistent estimates in the presence of correlation between a subset of the explanatory variables and unobserved provider effects, identify the effects of fixed provider characteristics. We apply this method, along with a range of other panel data methods, to national PROMs data for the financial years 2009 and 2010, in order to investigate whether ownership affects the quality of care provided to NHS patients undergoing hip replacement surgery.

We were concerned with endogeneity bias due to unobserved systematic differences in patient casemix between ISTCs and NHS providers, not accounted for in our casemix adjustment. We used a range of specification tests to select the optimal model. These indicate a lack of correlation between the explanatory variables and the unobserved heterogeneity, indicating that that this type of endogeneity does not exist here. As a result, the Hausman-Taylor model was not optimal here.

The results from the preferred random effects specification indicate that the quality of care provided at ISTCs is significantly greater than that provided at NHS institutions. This result is at odds to the concerns of many within the NHS that ISTCs were providing lower quality care (Bardsley and Dixon, 2011). This result is consistent with other studies in the ISTC literature, which indicate that the quality of care is significantly greater at ISTCs for hip replacement patients (Browne et al., 2008; Chard and Kuczawski, 2011). In an apparent critique of the use of ISTCs, Dickson and Anders (2011) outline some of the possible reasons for this positive effect. They point to surgeons and anesthetists in ISTCs not being distracted by the needs of patients not on their operating lists, not having their operating list disrupted by emergency cases, and not being distracted by having to teach medical students and mentor junior staff, as advantages over NHS staff, and reasons for higher outcomes in ISTCs. Using data on all, and not only a subset, of providers, strong case-mix adjustment and a rich

covariate set, and our methodological approach strengthens our findings in comparison with those in the previous literature.

5.2 Applying the Hausman-Taylor model to patient-provider data

Although the specification tests did not reveal the Hausman-Taylor estimator as the preferred model, the patterns in the estimates between models deserve some attention. We find that although the estimates of the effects of patient-varying variables are consistent between the Hausman-Taylor model and other specifications, estimates of the effects of fixed provider characteristics are considerably higher in magnitude and are estimated much less precisely. This imprecision may be due to the way the Hausman-Taylor model is implemented in Stata. While the xthtaylor command in Stata, which we employ, allows for an unbalanced panel through adjusting the GLS transformation, it does not adapt the instrument set as suggested in either Gardner (1998) or Allin (2011). Although he doesn't elaborate, Gardner states that weighting the provider averages in the instrument set is required to elicit an accurate model. Allin carries out a simulation exercise and finds that the standard deviations of "estimates minus parameter values" for fixed variables produced by the xthtaylor command are considerably higher than those when the instrument set adjustment is carried out. However, given this paper is currently unpublished; the degree to which results can be trusted is limited. Given our panel is severely unbalanced here, further work should implement these adjustments to check for any considerable changes in both the magnitude and precision of Hausman-Taylor estimates.

More work is also needed in examining why the Hausman-Taylor estimate for the ISTC effect lies above that estimated from the random effects model, and why the Mundlak estimate lies below that of the random effects model.

5.3 Limitations

Our use of the national PROMs dataset reduces the risk of sample selection bias compared to previous work but does not eliminate it. We analyse only those patients for whom information is available for both the dependent variable and all explanatory variables. The national PROMs data suffers from two types of selection not dealt with here. The first is unit non-response, whereby patients fail to complete the post-operative questionnaire and thus post-operative outcomes are not observed. Heckman correction type models could be used to test and correct for bias from this type of selection. The second is item non-response, whereby patients, although completing both questionnaires, fail to complete a subset of questions which are needed for the explanatory variables. Multiple imputation techniques are available but we are not aware of any statistical package which allows estimation of the Hausman-Taylor model adjusted for using imputed data.

Here we measure quality of care for only a single procedure, hip replacement, using only a single health outcome, the OHS. Previous literature indicates the degree to which quality differs between ISTCs and NHS providers differs dependent on both the procedure and the outcome studied (Browne et al., 2008; Chard and Kuczawski, 2011; Pérotin et al., 2013). This work could be extended to cover patients undergoing knee replacement, hernia repair and varicose veins surgery, and could analyse less condition-specific measures of health such as the EQ-5D index, rates of complications, and measures of patient satisfaction.

The scope of the specification tests we employed for model selection could be extended. In a study investigating whether factors explaining expectations of nursing home entry are consistent with those affecting actual nursing home entry, Lindrooth et al. (2000) employ a more robust set of checks for the applicability of the Hausman-Taylor model. As well as testing the consistency of HT estimates through comparison with estimates from a fixed effects model, they also employ two tests for overidentification, outlined in Staiger and Stock (1994), which test the strength of the instruments. A Durbin-Wu-Hausman test, also outlined in Staiger and Stock (1994), is also employed. In our setting this would test that there is no correlation between the explanatory variables and the patient-varying error. Any future application should follow this more robust procedure.

The aim of this paper was to apply the Hausman-Taylor model in a setting common to many studies in health economics, were providers are observed over multiple patients. Unfortunately, in this particular application this model was not required. Despite this, we believe that the Hausman-Taylor model could be very beneficial in analyzing other research questions where a fixed provider characteristic is of main interest and endogeneity exits.

5.4 Recommendations for policy

Our results indicate that the quality of care provided to hip replacement patients at ISTCs is greater than that provided at NHS providers. Wave 1 ISTC contracts for treating NHS patients were set to end by 2011/12, with Phase 2 contacts expiring between 2011 and 2017 (King's Fund, 2009). The Department of Health has stated that central procurement of ISTCs has ceased. It was understood that upon expiry, any new contracts with ISTCs will have to be agreed with PCTs, with arrangements based around the 'standard NHS contract' (King's Fund, 2009). However, it is unclear how this will change with Clinical Commissioning Groups (CCGs) now responsible for commissioning services. Although patient choice has been extended, evidence from 2007 suggests only 45% of patients recalled being given a choice. As a result, the future of ISTCs depends both on actions of patients and the GPs responsible for making referrals. Our results for hip replacement indicate there is substantial benefit for patients in keeping ISTCs operational. However, evidence on quality in other areas of elective surgery and on more health outcomes is required before any concrete recommendations can be made.

6 Bibliography

Allin, 2011. The Hausman–Taylor (1981) estimator for panel data: a note on the unbalanced case.

Bardsley, M., Dixon, J., 2011. Quality of care in independent sector treatment centres. Bmj 343, d6936–d6936.

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 Serv. Res. 8, 78.

Chard, J., Kuczawski, M., 2011. Outcomes of elective surgery undertaken in independent sector treatment centres and NHS providers in England: audit of patient outcomes in surgery. BMJ Br. Med. … 343, d6404–d6404.

Charlson, M., Pompei, P., 1987. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J. chronic … 40, 373–383.

Dawson, J., Fitzpatrick, R., 1996. Questionnaire on the perceptions of patients about total hip replacement. J. Bone Jt. ….

Department for Communities and Local Government, 2011. English indices of deprivation 2010. London.

Department of Health, 2000. The NHS Plan: a plan for investment, a plan for reform., Vasa.

Department of Health, 2002a. Delivering the NHS plan: next steps on investment, next steps on reform.

Department of Health, 2002b. Growing capacity. Independent sector diagnosis and treatment centres. London.

Department of Health, 2008. Guidance on the routine collection of Patient Reported Outcome Measures (PROMs).

Department of Health and Independent Health Care Association, 2000. For the benefit of patients. A concordat with the private and voluntary health care provider sector. London.

Dickson, M., Anders, N., 2011. Independent sector treatment centres are a busted flush. BMJ 343, d7696.

Dr Foster Health, 2009. Hip and knee replacement summary [WWW Document]. URL http://www.drfosterhealth.co.uk/features/hip-and-knee-replacement.aspx (accessed 1.12.13).

Gardner, R., 1998. Unobservable individual effects in unbalanced panel data. Econ. Lett. 58, 39–42.

Hart, O., Shleifer, A., Vishny, R., 1997. The proper scope of government: theory and an application to prisons. Q. J. ….

Hausman, J., 1978. Specification tests in econometrics. Econom. J. Econom. Soc. 46, 1251–1271.

Hausman, J., Taylor, W., 1981. Panel data and unobservable individual effects. Econom. J. Econom. Soc. 49, 1377–1398.

Health Committee, 2006. Independent Sector Treatment Centres. London.

Healthcare Commision, 2007. Independent sector treatment centres: A review of the quality of care. London.

King's Fund, 2009. Briefing: Independent sector treatment centres. The King's Fund.

Lindrooth, R.C., Hoerger, T.J., Norton, E.C., 2000. Expectations among the elderly about nursing home entry. Health Serv. Res. 35, 1181–202.

London School of Hygiene and Tropical Medicine, 2007. Patient Reported Outcomes Measures (PROMs) in Elective Surgery.

Mason, A., Street, A., Verzulli, R., 2010. Private sector treatment centres are treating less complex patients than the NHS. J. R. Soc. Med. 103, 322–31.

Mundlak, Y., 1978. On the pooling of time series and cross section data. Econom. J. Econom. Soc. 46, 69–85.

National Centre for Health Outcomes Development, 2005. Preliminary overview report for schemes GSUPIC, OC123, LP4 and LP5.

Pérotin, V., Zamora, B., Reeves, R., Bartlett, W., Allen, P., 2013. Does hospital ownership affect patient experience? An investigation into public-private sector differences in England. J. Health Econ. 32, 633–46.

Staiger, D., Stock, J., 1994. Instrumental variables regression with weak instruments. Econometrica 65, 557–586.

Sussex, J., 2009. How Fair? Competition between independent and NHS providers to supply non-emergency hospital care to NHS patients in England 1–32.

7 Tables

Table 1: Descriptive statistics

	ISTCs (13.0%)		NHS providers (87.0%)	
	Mean	SD	Mean	SD
Age (years)	68.115	9.238	67.888	10.863
Male	0.406	.	0.418	.
IMD - income deprivation	0.104	0.075	0.121	0.090
Had a previous surgery (pre-op)	0.045	.	0.104	.
Having a revision surgery	0.027	.	0.082	.
Assisted with post-operative questionnaire	0.048	.	0.077	.
Self-reported disability (pre-op)	0.453	.	0.605	.
Charlson index	0.200	0.477	0.347	0.707
EQ-VAS (pre-op)	69.176	20.391	64.968	21.562
Non-white	0.007	.	0.016	.
Living arrangements (pre-op)				
Partner/spouse/family/friends	0.751	.	0.734	.
Alone	0.250	.	0.260	.
Nursing home, hospital or other long-term care home	0.001	.	0.002	.
Other	0.003	.	0.004	.
Length of symptoms				
<1 year	0.175	.	0.137	.
1-5 years	0.657	.	0.663	.
5-10 years	0.101	.	0.111	.
>10 years	0.067	.	0.089	.
General health (pre-op)				
Excellent	0.072	.	0.046	.
Very good	0.358	.	0.276	.
Good	0.441	.	0.440	.
Fair	0.118	.	0.200	.
Poor	0.012	.	0.039	.
EQ-5D Mobility (pre-op)				
No problems	0.092	.	0.064	.
Some problems	0.906	.	0.932	.
Severe problems	0.001	.	0.004	.
EQ-5D Self care (pre-op)				
No problems	0.528	.	0.442	.
Some problems	0.468	.	0.546	.
Severe problems	0.004	.	0.012	.
EQ-5D Usual Activities (pre-op)				
No problems	0.080	.	0.063	.
Some problems	0.787	.	0.742	.
Severe problems	0.133	.	0.195	.

*All statistics are to 3 d.p.

Table 1 (continued): Descriptive statistics

		ISTCs (13.0%)		NHS providers (87.0%)	
		Mean	SD	Mean	SD
EQ-5D Pain/Discomfort (pre-op)					
	No problems	0.012	·	0.001	·
	Some problems	0.670	·	0.575	·
	Severe problems	0.318	·	0.415	·
EQ-5D Anxiety/Depression (pre-op)					
	No problems	0.653	·	0.580	·
	Some problems	0.320	·	0.373	·
	Severe problems	0.027	·	0.047	·
Comorbidities (pre-op)					
	Heart Disease	0.059	.	0.109	.
	High Blood Pressure	0.383	.	0.409	.
	Stroke	0.007	.	0.015	.
	Circulation problems	0.041	.	0.072	.
	Lung Disease	0.047	.	0.067	.
	Liver disease	0.004	.	0.005	.
	Kidney Disease	0.009	.	0.018	.
	Disease of the Nervous System	0.006	.	0.009	.
	Cancer	0.041	.	0.050	.
	Depression	0.053	.	0.070	.
	Arthritis	0.747	.	0.725	.
	Diabetes	0.060	.	0.093	.
Provider characterisitcs					
	Total volume	6125.104	3954.144	216489.540	126153.570
	PROMs volume	282.571	324.684	376.955	263.109
	HRG Count	134.387	57.556	955.237	203.305
Oxford Hip Score (OHS)					
	Pre-op	20.247	8.206	18.186	8.449
	Post-op	40.743	7.902	37.615	9.660
	Change	20.496	9.710	19.429	10.353
Observations		4,221		28,242	

*All statistics are to 3 d.p.

Table 2: Specification Tests

Test Name	Distribution	DF	Test Statistic	P-Value	Implications
Breusch-Pagan, OLS versus error component	χ^2	1	135.97	0.000	Significant variation in errors across providers
Hausman Test - random effects versus fixed effects	χ^2	44	44.89	0.434	Random effect is efficient and consistent
Hausman-Taylor Test - HT versus fixed effects	χ^2	44	2.50	1.000	HT is consistent, instruments valid

Table 3: Effect of provider type on OHS improvement: All models.

	OLS		Random Effects		Fixed Effects		Mundlak		Hausman-Taylor	
	Coeff	**Std. err**	**Coeff**	**Std. err**	**Coeff**	**Std. err**	**Coeff**	**Std. err**	**Coeff**	**Std. err**
ISTC	1.446***	(0.27)	1.651**	(0.62)	.	.	1.174	(0.66)	5.076	(3.38)
Pre-op OHS	0.190***	(0.01)	0.189***	(0.01)	0.187***	(0.01)	0.188***	(0.01)	0.188***	(0.01)
EQ-5D Mobility (Base=no problems)										
Some problems	0.622**	(0.20)	0.627**	(0.20)	0.612**	(0.20)	0.612**	(0.20)	0.613**	(0.20)
Severe problems	-1.283	(0.81)	-1.334	(0.81)	-1.369	(0.81)	-1.361	(0.81)	-1.361	(0.81)
EQ-5D Self care (Base=no problems)										
Some problems	-0.868***	(0.11)	-0.871***	(0.11)	-0.872***	(0.11)	-0.870***	(0.11)	-0.871***	(0.11)
Severe problems	-3.786***	(0.46)	-3.779***	(0.46)	-3.734***	(0.47)	-3.745***	(0.46)	-3.741***	(0.46)
EQ-5D Usual Activities (Base=no problems)										
Some problems	0.726***	(0.21)	0.748***	(0.21)	0.752***	(0.21)	0.757***	(0.21)	0.750***	(0.21)
Severe problems	0.146	(0.25)	0.185	(0.25)	0.190	(0.25)	0.196	(0.25)	0.189	(0.25)
EQ-5D Pain/Discomfort (Base=no problems)										
Some problems	-0.242	(0.47)	-0.285	(0.47)	-0.298	(0.47)	-0.292	(0.47)	-0.298	(0.47)
Severe problems	-0.412	(0.49)	-0.454	(0.49)	-0.473	(0.49)	-0.467	(0.49)	-0.472	(0.49)
EQ-5D Anxiety/Depression (Base=no problems)										
Some problems	-0.734***	(0.11)	-0.728***	(0.10)	-0.730***	(0.11)	-0.733***	(0.11)	-0.729***	(0.10)
Severe problems	-1.649***	(0.25)	-1.604***	(0.25)	-1.571***	(0.25)	-1.572***	(0.25)	-1.573***	(0.25)
General health (Base=excellent)										
Very good	-0.892***	(0.22)	-0.881***	(0.22)	-0.851***	(0.22)	-0.862***	(0.22)	-0.854***	(0.22)
Good	-2.206***	(0.22)	-2.197***	(0.22)	-2.166***	(0.22)	-2.179***	(0.22)	-2.169***	(0.22)
Fair	-4.336***	(0.25)	-4.340***	(0.25)	-4.324***	(0.25)	-4.338***	(0.25)	-4.325***	(0.25)
Poor	-6.188***	(0.35)	-6.186***	(0.35)	-6.137***	(0.35)	-6.151***	(0.35)	-6.143***	(0.35)
EQ-5D health scale (EQ-VAS)	-0.002	(0.00)	-0.002	(0.00)	-0.002	(0.00)	-0.002	(0.00)	-0.002	(0.00)
Self-reported disability (Base=no)	-1.989***	(0.10)	-1.941***	(0.10)	-1.932***	(0.10)	-1.928***	(0.10)	-1.932***	(0.10)
Length of Symptoms (Base=<1 year)										
1-5 years	-0.602***	(0.13)	-0.591***	(0.13)	-0.573***	(0.13)	-0.577***	(0.13)	-0.574***	(0.13)
5-10 years	-1.157***	(0.19)	-1.130***	(0.18)	-1.106***	(0.19)	-1.122***	(0.18)	-1.105***	(0.18)
>10 years	-1.693***	(0.20)	-1.682***	(0.20)	-1.674***	(0.20)	-1.675***	(0.20)	-1.673***	(0.20)
Previous Surgery (Base=no)	-1.904***	(0.27)	-1.931***	(0.27)	-1.937***	(0.27)	-1.932***	(0.27)	-1.937***	(0.27)
Self-reported comorbidities										
Heart disease	-0.394*	(0.16)	-0.368*	(0.16)	-0.367*	(0.16)	-0.361*	(0.16)	-0.366*	(0.16)
High blood pressure	0.086	(0.10)	0.083	(0.10)	0.077	(0.10)	0.079	(0.10)	0.080	(0.10)
Stroke	-0.546	(0.39)	-0.489	(0.39)	-0.474	(0.39)	-0.457	(0.39)	-0.473	(0.39)
Circulation problems	-2.394***	(0.19)	-2.362***	(0.19)	-2.352***	(0.19)	-2.348***	(0.19)	-2.357***	(0.19)
Lung disease	-0.082	(0.20)	-0.062	(0.19)	-0.048	(0.19)	-0.050	(0.19)	-0.050	(0.19)
Diabetes	-0.579**	(0.18)	-0.568**	(0.18)	-0.569**	(0.18)	-0.561**	(0.18)	-0.569**	(0.18)
Kidney disease	-0.073	(0.36)	-0.137	(0.36)	-0.148	(0.36)	-0.146	(0.36)	-0.146	(0.36)
Nervous system	-1.068*	(0.50)	-1.155*	(0.50)	-1.222*	(0.50)	-1.228*	(0.50)	-1.202*	(0.50)
Liver disease	-1.236	(0.64)	-1.209	(0.63)	-1.220	(0.64)	-1.221	(0.63)	-1.218	(0.63)
Cancer	0.529*	(0.21)	0.538*	(0.21)	0.545*	(0.21)	0.548*	(0.21)	0.543*	(0.21)
Depression	-1.720***	(0.19)	-1.718***	(0.19)	-1.729***	(0.19)	-1.725***	(0.19)	-1.728***	(0.19)
Arthritis	-0.491***	(0.11)	-0.467***	(0.11)	-0.457***	(0.11)	-0.460***	(0.11)	-0.457***	(0.11)
Observations	32463		32463		32463		32463		32463	

*p<0.05, **p<0.01, ***p<0.001

Table 3 (continued): Effect of provider type on OHS improvement. All models.

	OLS		Random Effects		Fixed Effects		Mundlak		Hausman-Taylor	
	Coeff	**Std. err**	**Coeff**	**Std. err**	**Coeff**	**Std. err**	**Coeff**	**Std. err**	**Coeff**	**Std. err**
Gender (Base=female)	0.908***	(0.10)	0.921***	(0.10)	0.927***	(0.10)	0.922***	(0.10)	0.927***	(0.10)
Non-white (Base=white)	-2.438***	(0.38)	-2.525***	(0.39)	-2.480***	(0.39)	-2.495***	(0.39)	-2.494***	(0.39)
Age										
Level	-0.032***	(0.01)	-0.035***	(0.01)	-0.036***	(0.01)	-0.036***	(0.01)	-0.035***	(0.01)
Power 2	-0.004***	(0.00)	-0.003***	(0.00)	-0.003***	(0.00)	-0.003***	(0.00)	-0.003***	(0.00)
Power 3	-0.000	(0.00)	-0.000	(0.00)	-0.000	(0.00)	-0.000	(0.00)	-0.000	(0.00)
Power 4	0.000**	(0.00)	0.000**	(0.00)	0.000**	(0.00)	0.000**	(0.00)	0.000**	(0.00)
Power 5	0.000*	(0.00)	0.000*	(0.00)	0.000*	(0.00)	0.000*	(0.00)	0.000*	(0.00)
Living arrangements										
Alone	-0.119	(0.11)	-0.126	(0.11)	-0.128	(0.11)	-0.128	(0.11)	-0.128	(0.11)
Care home[a]	-0.929	(1.19)	-0.856	(1.18)	-0.847	(1.19)	-0.809	(1.19)	-0.848	(1.18)
Other	-0.763	(0.74)	-0.744	(0.74)	-0.726	(0.74)	-0.730	(0.74)	-0.729	(0.73)
Post-op assistance (Base=no)	-2.088***	(0.18)	-2.062***	(0.18)	-2.040***	(0.18)	-2.045***	(0.18)	-2.046***	(0.18)
IMD income score	-8.390***	(0.54)	-7.412***	(0.56)	-6.953***	(0.57)	-6.965***	(0.57)	-7.028***	(0.56)
Charlson index	-0.130	(0.08)	-0.117	(0.08)	-0.119	(0.08)	-0.116	(0.08)	-0.119	(0.08)
Revision surgery (Base=no)	-3.661***	(0.31)	-3.604***	(0.31)	-3.595***	(0.31)	-3.597***	(0.31)	-3.595***	(0.31)
HRG count	0.000	(0.00)	0.001	(0.00)	.	.	0.001	(0.00)	0.005	(0.00)
Volume: PROMs procedures	0.001***	(0.00)	0.002***	(0.00)	.	.	0.001*	(0.00)	0.003*	(0.00)
Volume: All procedures	-0.000	(0.00)	-0.000	(0.00)	.	.	0.000	(0.00)	-0.000	(0.00)
Constant	40.205***	(0.74)	39.597***	(0.97)	40.897***	(0.67)	49.070***	(12.45)	35.753***	(3.68)
Observations	32463		32463		32463		32463		32463	

*p<0.05, **p<0.01, ***p<0.001

[a]Nursing home, hospital or other long-term care home