Do financial incentives trump clinical guidance? Hip replacement in England and Scotland

Do Financial Incentives trump Clinical Guidance? Hip Replacement in England and Scotland

Irene Papanicolas
***corresponding author**
Department of Social Policy
London School of Economics
Houghton Street
London
WC2A 2AE
i.n.papanicolas@lse.ac.uk

Alistair McGuire
Department of Social Policy
London School of Economics
Houghton Street
London
WC2A 2AE

Abstract
Following devolution in 1999 England and Scotland's National Health Services diverged, resulting in major differences in hospital payment. England introduced a case payment mechanism from 2003/4, while Scotland continued to pay through global budgets. We investigate the impact this change had on activity for Hip Replacement. We examine the differential financial reimbursement attached for uncemented Hip Replacement in England, which has been more generous than for its cemented counterpart, although clinical guidance from the National Institute for Clinical Excellence recommends the later. In Scotland this financial incentive does not exist. We use a difference-in-difference estimator, using Scotland as a control, to test whether the change in reimbursement across the two countries had an influence on treatment. Our results indicate that financial incentives are directly linked to the faster uptake of the more expensive, uncemented Hip Replacement in England, which ran against the clinical guidance.

Keywords: Financial Incentives; DRGs; Hip Replacement; Activity based payment

1. Introduction

Many health care systems are using competition within managed care environments, where for example hospitals face fixed price regulation, to handle the trade-off between cost and delivery in quality of care. While there is general agreement that the accompanying payment systems adopted to encourage competition do affect provider performance, empirical evidence to support this view remains relatively sparse. The empirical evidence that does exist largely draws on US data relating to the introduction by Medicare of prospective DRG payment to hospitals in the mid-1980s, and even here few studies consider reactions to subsequent changes in fixed prices (see, for example Cutler, 1995; Gilman, 2000; Dafny, 2005)[1]. Moreno-Serra and Wagstaff (2010) provide examples of the literature outside of the USA, as well as evidence on system-wide effects of payment reform in Europe and Asia. Of this literature few have considered the impact of price increases on activity once a fixed system is in place (Dafny, 2005; White and Yee, 2013; He and Mellor, 2013). While there has been analysis of payment incentive effects in the UK, once again the empirical literature relating this to changes in activity is limited[2].

The relative lack of empirical evidence relating to the UK hospital sector and the introduction and operation of fixed payments is surprising given the extensive reforms that have been underway in the UK since the mid-1990s. In NHS England, part of the UK National Health Service (NHS), the introduction of competition amongst hospitals around the mid-2000s has been argued to promote efficiency and improve quality of outcome within the health care sector and has been supported by empirical evidence provided by Cooper et al (2011) and Gaynor et al (2013). These findings are in line with a growing literature on competition and case-based payment systems (see Gaynor, 2012 for a review)[3]. Such competition has in fact been accompanied by increased regulation, partly to guarantee that clinical standards are maintained despite competition for funds. In particular national clinical guidelines, as specified by the National Institute for Health and Clinical Excellence (NICE), form the basis of managing health care within the English NHS. These guidelines cover a wide range of interventions and are based on assessments of the clinical evidence in specific areas to help to ensure that providers are maintaining, or even improving quality standards in the delivery of the care across specific disease areas.

With NICE already in existence, the English NHS introduced case-based payment system in 2003/4, where they linked individual case groupings – or Health Related Groups (HRGs)[4] – to specific reimbursement rates derived from treatment costs. This case-based payment system is essentially a form of Diagnostic Related Group (DRG) reimbursement, and is

[1] An interesting distinction between the effect of incentive changes on the marginal reimbursement effect compared to the average reimbursement effect is made by Cutler (1995), who maps the incentive effects in a move from Medicare cost-based reimbursement to DRG reimbursement. This is similar to the change in the UK from contract-based reimbursement to Payment by Results (PbR) reimbursement, however the UK contract-based reimbursement prior to the introduction of PbR was largely based on block contracts covering populations rather than reimbursement of the volume of care undertaken. This of course renders the analogy to marginal and average reimbursement redundant within the UK setting.

[2] The literature on related topics within the UK covers, for example, how competition and payment by results has affected outcomes (see Gaynor et al, 2012 and Propper, 2012 for reviews of this literature), how it has affected the mix of hospital activity (Farrer et al, 2009); how regulation has affected waiting times (Propper et al, 2010); GP doctor behaviour and payment by performance (Gravelle et al, 2010). See Busse et al (2011) for discussion of the literature on hospital payment systems in Europe.

[3] Although see Gravelle et al (2012) for a wider view.

[4] Further specific information on HRGs can be found at (Street and Dawson, 2002; Mason et al. 2011)

referred to as Payment by Results (PbR). The PbR reimbursements are nationally agreed tariffs, set by the Department of Health and used in England by purchasers of health care to reimburse individual providers -mainly hospitals - for the provision of treatment. The tying of interventions to specific levels of reimbursement provides a means of testing the importance of financial consideration in a managed care environment, particularly if clinical guidance exists within that specific disease area.

Of particular interest is the case of Hip Replacement, an extremely common procedure with substitutable treatment options available and where, at least in other systems, patient preferences and financial incentives have been shown to play a role in treatment up-take (Doiron, Fiebig and Suziedelyte, 2014). Hip Replacement presents a unique opportunity to study the incentives created by English PbR payment system. In a Hip Replacement, two main types of prostheses are available; cemented and uncemented. Both types have been around since the 1970s and clinical evidence suggests that both prostheses have comparable rates of success (Abdulkarim et al, 2013). Until recently, the vast majority of Hip Replacements performed in the UK used cemented prostheses, although the number of uncemented Hip Replacements undertaken has increased substantially in the past decade. This change in prosthesis use has coincided with the introduction of separate reimbursement rates for the two types of prostheses, which provide a more generous surplus for the uncemented implant, possibly to cover the longer operating times required to fit the uncemented device. The increase in up-take of the more generously reimbursed implant is in spite of recommendations from the NICE that favour the use of cemented prostheses in Hip Replacements (NICE, 2000; updated in 2013). While NICE guidance and quality standards are not absolutely mandatory, they are used by NHS regulators to establish acceptable levels of care, and if required health care providers must defend any individual treatment decisions which run contrary to NICE guidance.

Hip replacements are also of interest because individual hospital providers control the procurement practices with respect to prostheses; thus managers have potentially more influence over the type of procedure finally implemented than in other cases. In their analysis of procurement practices in the NHS, Davies and Lorgelly (2013), focused on a case-study of Hip Replacement and the purchasing of hip prosthesis. They note that in the UK NHS, the hospital through its centralised procurement policies – as opposed to the individual surgeon – determines the specific prostheses to purchase and negotiates quantities and prices with the suppliers. Individual surgeons feed their preferences into the procurement process; acting as an agent for their patients by including patient characteristics within their own surgical preferences. The particular prostheses purchased at the hospital level thus reflect individual surgeon preferences, historical procurement practices, prices and reimbursement levels. Davies and Lorgelly (2013), also note that, if volume discounts are available this may lead to specialisation in prosthesis type. In other words, characteristics of hospital behaviour, as informed by surgical assessment, will determine the specific prosthesis to be purchased by any hospital, at any point in time. There will inevitably be a trade-off, at the hospital level, between management and surgical preferences. However, it is the ability to centralise procurement decisions and to hold stock that provides a mechanism through which hospitals can control the type of device, and therefore the revenue generated from this relatively common procedure.

With regards to Hip Replacement prosthesis, we have then a situation in the UK NHS, where England has different fixed DRG-type reimbursement rates (PbR) for two common, substitutable procedures - cemented and uncemented replacements – while at the same time,

in England NICE recommends the less expensive cemented replacements above uncemented replacements in their clinical guidance. In Scotland, as providers are not reimbursed for cases treated, no such financial incentive exists to influence choice. This situation provides a means of analysing, in a controlled manner, the impact financial incentives can have on specific procedure up-take at the individual hospital level, for a procedure where prosthesis type does not affect clinical outcome. This provides a unique case-study of individual hospital purchasing decisions, made through managed procurement practices, where decisions may be influenced by revenue generation given that prosthesis type has no influence on patient outcomes.

The 1999 devolution has presented a natural experiment in health care provision within an NHS system as England and Scotland have diverged substantially in the reforms they have implemented to meet their National Health Service objectives –essentially creating two different NHS systems within the UK (Leys, 1999; Pollock, 1999). The English NHS has embraced market mechanisms and cooperation with the private sector, while the Scottish NHS has moved in the opposite direction, and created a highly centralised system that maintains trust in its providers to allocate resources effectively, and strives for improvement through integration (Steele and Cylus, 2014; Greer, 2006).

One of the main differences in health policy that has emerged in the years following devolution has been in the funding of inpatient hospital care. Prior to 1997, England and Scotland funded inpatient care in broadly the same way: health care purchasers and providers negotiated the services that would be provided through bulk contracts (Ham, 2004). Scotland has moved away from this funding system and since 2004 has funded inpatient care through the allocation of global hospital budgets (Scottish Parliament, 2004). England on the other hand has further supported the internal market by moving away from the bulk contract system of funding hospital episodes to a fix-priced activity-based payment system, of DRG-type reimbursement, known as Payment by Results (PbR), introduced in in 2003/04.

Given the divergence in funding for inpatient activity across the two nations, we use Scottish NHS hospitals as a control group within a difference-in-difference style estimator, as well as employing a large number of robustness checks, to test whether the up-take of the more expensive uncemented prosthesis in England was influenced by reimbursement levels, at a time when the less expensive cemented prosthesis was being recommended by NICE. Our results add to the literature on the impact of financial incentives on individual providers in a managed care setting by providing a specific example. Our conclusions suggest that English NHS hospitals did indeed have higher up-take rates of the more generously reimbursed uncemented Hip Replacements than the (Scottish) control group providers after PbR had been introduced in England, despite the English clinical guidance recommending cemented Hip Replacements. The ability to hold stock and for hospitals to manage procurement when acquiring prostheses allows individual in-patient providers of Hip Replacement flexibility in their purchasing decision.

The precise question we examine in this paper is: To what extent any observed change in the selection of Hip Replacement prosthesis type in England, can be related to the financial incentive introduced by the differential reimbursement of treatment interventions by DRG/PbR payment? As we document below, in the case of Hip Replacement, the financial incentive associated with PbR reimbursement provided a signal opposite to that outlined by the NICE technology appraisal. As such, the examination of this question provides a unique

study of the impact that financial incentives have on clinical practice in an environment which is highly regulated, but where individual hospitals have discretion over procurement.

The paper proceeds by briefly outlining the funding differences that exist in England and Scotland, as well as highlighting the pattern of prosthesis up-take in both countries over a 10-year period, which motivates the use of Scottish providers as a control. The following sections outline the data and methods, after which we present results and offer conclusions.

2. Background

This study is concerned with the impact that the different incentives introduced for the provision of care across England and Scotland had on the uptake of competing procedures; cemented and uncemented Hip Replacement. Given the timeframe we are particularly interested in examining the role financial incentives can play in influencing clinical practice. As previously mentioned, while cemented and uncemented Hip Replacement represent different procedures both are used to treat similar patients. Cemented replacement is the older technique (developed about 40 years ago), and is assumed to have a shorter and easier recovery period, while uncemented replacement, developed 20 years ago, avoids the possibility of loosening parts and the breaking off of cement particles. To date, clinical evidence suggests that both have comparable rates of clinical success (Abdulkarim et al, 2013).

Over the past decade different types of incentives have been introduced that favour one implant over the other, as demonstrated for the two countries in Figure 1.

< Insert Figure 1 here >

Immediately after Devolution, in 2000 NICE, which is based in England, conducted a technology appraisal (TA2) that recommended that cemented Hip Replacements should be performed over cementless procedures (NICE, 2000), this was updated in 2013 and the same guidance was reconfirmed. This guidance is based on the difference in cost between the two prostheses (cemented being cheaper on average) as there was little long term (10 years or more) evidence to suggest which implants had better outcomes (in particular, revision rates). However, with the introduction of PbR in England in 2003/4 financial incentives were introduced that favoured the uncemented procedures over their cemented counterpart. As of 2003/4, PbR was introduced in England, where reimbursement became based on actual activity undertaken, defined by Health Related Groups (HRGs)[5], following a national HRG tariff.

PbR was phased in gradually over the years 2003/4 to 2006/7. It was initially introduced for 15 HRGs in its first year of implementation, extended to 33 HRGs in the second year, and then phased out to cover Foundation Trusts (self-managing NHS hospitals) initially and then for all providers of elective and emergency inpatient stays during the following years. Hip Replacement, coded as HGR H02, was one of the first 33 HRG groups to be allocated a reimbursement level in the phasing out of the PbR policy. After one year, this HRG group

[5] While HRGs were initially introduced in the late 1990s as an exercise to help define hospital activity costs based on clinically meaningful clinical groupings (Street and Dawson, 2002), it is only under PbR that they became attached to reimbursement.

(H02) was split into two separate HRGs (H80 and H81), representing cemented and uncemented primary Hip Replacement respectively, thus moving towards a differential reimbursement for each of the two procedures from 2005 onwards. This differential reimbursement continued until a new grouper was introduced in 2009, and the categorization of HRGs stopped reflecting the type on implant and instead reflected the severity of the patient. Hip Replacement reimbursement has further changed since the introduction of Best Practice Tariffs for Hip Fracture in 2010, and these do not differentiate between cemented and uncemented prostheses either[6].

The price of the national HRG tariff, updated yearly, is determined by the average of the costs calculated by all hospitals for each of their HRGs which includes labour, equipment and hospital costs, with small cost adjustments for geographical variation, labour market conditions and excessive length of stay for appropriate reason. There is a three-year delay[7] between hospitals submitting cost data and these data being converted into prices, and so an inflationary adjustment is also made to each HRG (Mason et al, 2011). HRG classifications themselves are updated and refined, through the continual updating of 'groupers'; the last of these updates took place in 2009 when HRG 4.0 was phased in.

Table 1 shows the tariff for the different HRG groupings over this period, alongside the corresponding average costs. In all years after the separation of HRG H02 into two distinct HRGs, the tariff for the elective cemented procedure (H80) has been reimbursed at a higher rate than the elective uncemented procedures, while the tariff for the non-elective procedures reimburses the uncemented procedure (H81) more highly. Although one might expect this difference to be linked to the underlying costs of these procedures, this does not appear to be the case, as illustrated by the national reference costs. These costs show the uncemented procedure to be cheaper in both settings[8]. In addition, the difference in the underlying cost and the tariff price is always larger for the uncemented procedures[9]. This suggests that greater financial surplus is made through performing the uncemented procedure rather than the cemented, particularly for emergency procedures.

< Insert Table 1 here >

An estimate of the average revenue gain, R_{jt}, to be made from switching to uncemented replacement from cemented replacement, for each admission type j, for a given year t, is presented in the final columns of Table 1; calculated as:

$$R_{jt} = \left(UT_{jt} - \overline{UC}_{jt-3}\right) - \left(CT_{jt} - \overline{CC}_{jt-3}\right)$$

[6] As Hip Replacement was one of the first 33 conditions introduced, it was phased out to all providers in 2004. Thus, all types of English hospitals (Acute Care Trusts, Foundation Trusts) received the same reimbursement. This is also true for the introduction of the differential tariff the next year. The only exception to this are the ISTC providers who were given a tariff uplift to presumably cover capital costs, however the information on the public uplift was not made publicly available.

[7] Since the introduction of the new HRG grouper in 2009 this has now been reduced to a two-year delay.

[8] While the uncemented prostheses are on average more expensive than the cemented ones, the surgery is faster to perform (approximately 15 min according to Yates et al (2006)) which may account for the lower average cost.

[9] Reference costs are not adjusted for age or other patient characteristics. In their conversion to the tariff they undergo data filtering, cleaning, spell conversation, adjustments to reflect tariff scope and structure and price adjustments. Also note that these reference costs will not include any discounts to purchasers.

where $\overline{UT_{jt}}$ and $\overline{CT_{jt}}$ represent the tariff for uncemented and cemented Hip Replacement by admission type, and $\overline{UC_{jt}}$ and $\overline{CC_{jt}}$ represent the average cost for the cemented and uncemented HRGs. As there was a three-year delay between hospitals submitting cost data and these data being converted into prices, the t-3 subscript is attached to the average costs.

These estimates indicate that in the case of non-elective Hip Replacement large revenue gains, approximately £900-£1700 per case, on average, can be made by switching to the uncemented implant over the period in question. The potential revenue gains are smaller in the case of the elective surgery, but are for the most part positive, ranging from an approximate loss of £8 in one specific year, to a gain of about £400 per case on average. The actual prosthesis prices are bound to influence individual hospital procurement practices where, as noted above hospital management and clinicians make explicit purchasing decisions on the type of prosthesis to stock. The national reference costs are used here, as the commercial prices of prostheses are not publicly available, to indicate potential financial surplus and are generally regarded as indicative of true treatment costs. These reference costs are publicly available as averages over all hospitals. The more efficient hospitals have the potential to make larger financial gains than those indicate above. Moreover, these estimates of potential gain do not factor in any implant discounts that can be made through negotiations between individual providers and suppliers.

In Scotland no differential financial incentives exist, and the technology assessments issued by NICE also do not automatically apply, as clinical guidance is provided by the Scottish Medicines Agency. The NICE guidance relating to Hip Replacement was in fact applied latterly in Scotland, (although not until 2014), where it was introduced by Healthcare Improvement Scotland (HIS). Scotland does however use HRGs to code activity, extending their use from their pre-devolution implementation. From 2005 onwards Scotland also started to calculate an HRG Tariff, which was to be used as a costing tool to promote efficiency. The derivation of the Scottish National tariff is based on the English National Tariff, although differences do exist. While HRGs and the Scottish National Tariff are used to inform service delivery, they are not used for reimbursement. Scotland therefore provides a useful control case to observe the choice of Hip Replacement in a part of the UK NHS were the financial incentives do not apply.

< Insert Figure 2 here >
< Insert Figure 3 here >

Figure 2 illustrates the levels of cemented and uncemented activity across England and Scotland for the duration of our sample (1996-2012). This figure clearly illustrates the differential uptake of uncemented procedures in England over the 2000s relative to Scotland. In particular, it appears that the levels of uncemented procedures in England begin to increase from about 2002/3, but only begin to coincide with a fall in uncemented activity from about 2004/5 - coinciding with the split of the financial incentive. While the Scottish trends remain relatively flat in comparison, it appears that from the period 2002/2006, Scotland experiences an increase in cemented activity, although from 2007 onwards uncemented procedures increase while cemented activity falls.

However, as the magnitude of activity in much larger in England, it is perhaps more informative to also examine the proportions of Hip Replacement activity over the same period. Figure 3 illustrates the proportion of uncemented cases to the total (uncemented +

cemented) over the time period being investigated. Up until 2003, the fraction of uncemented Hip Replacements performed in the two countries remains relatively stable; at about 10% in England and under 5% in Scotland. In England, and with the introduction of PbR in 2003, the proportion of uncemented procedures starts to rapidly increase, until it reaches is peak in 2010, making up 50% of the total. In both countries the increase in uncemented proportions coincides with the wider use of HRGs. In England 2003 is the introduction of PbR, while in Scotland 2005 marks the introduction of the use of HRGs as a costing/efficiency tool.

As Table 1 shows the national reference costs for uncemented Hip Replacement are lower for both Elective and Non-Elective procedures, thus providing hospitals in either country – operating under a budget or a fixed price reimbursement – to opt for the uncemented Hip Replacement if pursing efficiency gains by some financially astute providers. The financial incentive, introduced by split payments, which reimburse the cheaper Uncemented prosthesis at a higher rate, only magnify this incentive in England. In Scotland, uncemented procedures also increase but not until 2005, when the National Scottish Tariff was introduced for managerial - not reimbursement - reasons, and even then the rate of increase is less rapid than in England and, as shown in Figure 2, is associated with much smaller changes in levels of activity.

< Insert Figure 4 here >

Given that the choice of Hip Replacement procedure is likely to be closely tied with the procurement practices of prostheses we are also interested in examining the proportion of cemented and uncemented Hip Replacements undertaken in each hospital, to understand if the switch to uncemented is driven by particular hospitals switching all their prosthesis or by a selection of the uncemented procedure for particular types of patients across hospitals. Figure 4 illustrates the frequency of uncemented Hip Replacement cases to total (uncemented + cemented) by provider across England and Scotland for selected years. The first panel shows the frequencies in 1996, the first year in our data. In this panel, the histogram illustrates that most providers, in both countries, are not performing any uncemented Hip Replacements, and those who are, are for the most part performing it in less than 25% of their cases. The second panel, shows the situation in 2005, a couple years after PbR has been rolled in. While the situation in Scotland is largely unchanged, in England more providers are performing some uncemented procedures, with a few providers providing almost entirely uncemented procedures. In 2009, the last year of the differential incentive in England, we see that the situation in England is split such that about 50% of providers are performing uncemented procedures on more than half the of their patients. In Scotland more providers have started to provide uncemented Hip Replacements – a few providing almost exclusively uncemented procedures - but the majority continue to provide mostly or exclusively cemented Hip Replacements. Finally, in 2012 – the last year of our data, England remains split, such that about half the providers are providing mostly uncemented procedures, and half mostly cemented, while in Scotland more providers are providing uncemented procedures but cemented still dominates.

It is important to note that as activity changes over time as shown by these Figures, with more providers procuring more uncemented prostheses, any financial incentive does not give rise to complete substitution. This no doubt partly reflects existing clinical practices and preferences, as well as the suitability of prosthesis type for individual patients. Any potential financial incentive is therefore mediated by clinical practice. There is switching on average across all providers, indicating a common incentive, but the switching is not complete.

3. Data

The data used to conduct this analysis are drawn from two administrative databases; the English Hospital Episode Statistics (HES) and the Information Services Division (ISD) of NHS Scotland. Both data sets contain records for all NHS patients admitted to all NHS hospitals in each country, with information on all medical and surgical specialties performed. The data also provide information on patient characteristic data (e.g. age), clinical information (e.g. diagnoses using ICD-10 codes, procedures using OPCS codes and HRG codes), mode of admission and details where the patient was treated. The HRG codes used in the ISD data are calculated based on the methodology used in England.

All individual cases coded with HRG 3.5 grouper codes, H80 and H81 or OPCS codes W371 or W381 for Hip Replacement were extracted for the years 1996-2010, as after this date the same (Best Practice Tariffs) were applied to all Hip Replacement procedures in England. Where a different HRG grouper was used, the HRG 3.5 grouper was applied to the data to allow comparisons across the time period. The change in groupers over the time periods makes it difficult to use them over the period being investigated, and so we prefer the surgical OPCS 4 codes to identify uncemented (code W381) and cemented (code W371) procedures which ensure better consistency. The number of cases for each of the HRG groups and OPCS categories were aggregated for each hospital, separately for each year of the sample, and exported into a newly constructed panel, together with aggregated statistics on the patients treated in each hospital, and hospital status information. Table 2 presents descriptive statistics on variables of interest.

< Insert Table 2 here >

We use the proportion of uncemented Hip Replacement to total (cemented + uncemented) procedures as the dependent variable in most of our specifications. The proportions variable allows us to capture relative change across the two Hip Replacement techniques, thus directly incorporating any potential substitution from cemented to uncemented Hip Replacements over the time period. It provides, we believe, a stronger test of change is activity than any observed change in levels.

The aggregated statistics used as controls for severity and patient characteristics are constructed as the mean values for each hospital, for the cases being investigated, derived from individual patient level data. These include age, sex and severity, measured by the Charlson co-morbidity index. The Charlson co-morbidity index controls for a total of 22 conditions[10], and is constructed by assigning a score to each co-morbid condition depending on the one-year risk of death associated with it, and summing these scores up (Charlson et al., 1987). Finally, we also construct a measure of volume for each hospital, which simply measures the total number of Hip Replacements (both cemented and uncemented) undertaken at each hospital for every year of the data.

[10] These are Myocardial Infarction, Congestive Cardiac Failure, Peripheral Vascular disease, Dementia, Cerebrovascular disease, Chronic Lung disease, Connective Tissue disease, Ulcer, Chronic Liver disease, Hemiplegia, moderate or severe Kidney disease, Diabetes, Diabetes with complications, Tumor, Leukemia, Lymphoma, moderate or severe Liver disease, Malignant Tumor, Metastasis and AIDS.

The aggregated dataset at the hospital level, described above, is then used to examine the differences in hospital activity across the two country settings and attributed to changes in English financial reimbursement rates.

4. Methods

We use regression analysis based on the proportion of uncemented Hip Replacements to test whether hospitals in the PbR environment had a higher uptake of these procedures, relative to hospitals in Scotland as they did not face changes in financial incentive and did not have to enforce the NICE clinical guidance. We use a difference-in-difference (DiD) approach, as specified below:

$$\left(\frac{Uncemented}{Total}\right)_{ht} = \alpha + \beta country + \gamma PbR + \delta(PbR_t * country) + \lambda\,(year\ dummy_t) + \psi' controls_{ht} + \varepsilon_{ht} \quad (1)$$

In this, our preferred empirical specification, the dependent variable is the proportion of uncemented Hip Replacements to total (cemented + uncemented) Hip Replacement procedures at hospital *h*, during year *t.* The variable *PbR* is a dummy set equal to 1 from the year 2005 (the year PbR payment is attached to the separate Hip Replacement procedures in England). We interact this variable with our country identifier (*country:* England being the treatment country = 1; Scotland as control = 0). Finally, we include year dummies and a further number of controls relating to hospital (hospital type and volumes of Hip Replacement activity) and aggregate patient characteristics (average age, gender, co-mobridity), as defined in the data section above.

Our coefficient of interest is δ, as it captures the difference in the average proportional increase of uncemented Hip prostheses in each hospital in England before and after the introduction of PbR in England in 2005 minus the uptake in the average Scottish hospital (the control group), before and after the introduction of PbR in England. Holding all else constant the intercept, (α), captures the average proportion of uncemented hip prostheses in Scotland prior to 2005. Then, following standard DiD interpretation. The sum $\alpha + \beta$ captures the uptake in uncemented prostheses in Scotland post 2005. Finally, the sum $\alpha + \beta + \gamma + \delta$ captures the differential uptake in England compared to our control population after 2004. The inclusion of the time dummies provide a stronger test of the difference-in-difference estimate, δ.

We test our difference-in-difference model, by running a difference in trends estimator to examine the different phases of policy identification. We adopt this differences-in-trends specification to take account of the manner of hospital procurement in Hip Replacement, as we expect individual hospitals to run down their stock and replace stock, in line with the change in financial incentives after the introduction of the PbR, through procurement changes. In other words, the policy impact we assume will roll-out over time and will not necessarily be a distinct one-off change in activity. In addition, we expect that once the financial incentive is removed (2009-2012), with the introduction of the new HRG grouper, and the Best Practice Tariffs, this procurement process may change as they become accustomed to the new financial incentives. This model is specified below:

$$Uncemented/Total_{ht} = \alpha + \beta country + \gamma T + \delta PbR\{t - \tilde{t_1} | t \geq \tilde{t}_1\} + \zeta PbR\{t - \tilde{t}_2 | t \geq \tilde{t}_2\} + \eta country * PbR\{t - \tilde{t_1} | t \geq \tilde{t}_1\} + \theta country * PbR\{t - \tilde{t_2} | t \geq \tilde{t}_2\} + \varphi year + \lambda(year\ dummy_t) + \quad (2)$$

$\psi' controls_{ht} + \varepsilon_{ht}$

The dependent variable is the proportion of uncemented Hip Replacements to total (cemented + uncemented), as in equation (1). Our yearly spline, represented by $PbR\{t - \tilde{t} | t \geq \tilde{t}\}$ where *t* is a running counter of the years since 1996, where $\tilde{t}_1$ is the first break point in the spline, is set at the year 2005 (the year PbR payment is attached to the seperate Hip Replacement procedures in England), and $\tilde{t}_2$ is the break point in the spline again at 2009 when the incentive is removed. We interact these variables with our country identifier (*country:* England being the treatment country = 1; Scotland as control = 0). The variable *T* is a time trend to capture any systematic changes in practice over time, and as with the other models we include a further number of controls relating to hospital and patient characteristics and year dummies.

Imposing restrictions $\gamma = \varphi = 0$ returns a standard difference-in-trends estimator. More generally, η and θ, are the coefficients of interest as they are the spline-based difference-in-trends estimators that captures the difference in the average uptake of uncemented Hip prostheses in each hospital in England before and after the introduction of PbR in England in 2005 (and after its removal in 2009) minus the uptake in the average Scottish hospital (the control group). Holding all else constant the intercept, (α), captures the average volume of uncemented hip prostheses in Scotland prior to 2005. Then, following standard DiD interpretation $\alpha + \beta$ captures the incremental uptake in uncemented prostheses in England prior to 2005. The sum $\alpha + \delta$ captures the uptake in uncemented prostheses in Scotland post 2005, while $\alpha + \zeta$ captures the uptake in uncemented prostheses in Scotland post 2009. Finally, the sum $\alpha + \beta + \delta + \eta$ captures the differential uptake in England compared to our control population after 2005, and the sum $\alpha + \beta + \zeta + \theta$ captures the differential uptake after 2009. The inclusion of the time trend and the time trend country interaction provides a stronger test of the difference-in-trends estimates, η and θ.

As a final test on our basic specification we run an additional specification that allows us to make better use of the individual patient level data. In particular, we are interested in controlling for any individual patient characteristics which may increase the likelihood of receiving a cemented or uncemented Hip Replacement. To do this we estimate the hospital-specific effect on uncemented activity, θ, from a patient-level equation of the form:

$$Uncemented_{iht} = \theta_{ht} + \gamma' controls_{iht} + \varepsilon_{iht} \quad (3)$$

where the dependent variable is uncemented Hip Replacement, X is a set of patient characteristics (age, sex, co-morbidity, type of admission), ε is the error term, and *i* indexes the individual patient. The model is estimated separately for each year of the data in our sample (1996-2012) and the estimate, θ, is extracted. We then use θ as the dependent variable in our difference in difference model as specified above, in equation (3).

$$\theta_{ht} = \alpha + \beta country + \gamma PbR + \delta(PbR_t * country) + \lambda\,(year\ dummy_t) + \psi' controls_{ht} + \varepsilon_{ht} \quad (4)$$

All models are run with standard errors clustered at the hospital level and with random effects. The hospitals, *h,* are split into the treatment (England) and control (Scotland) groups, foregoing the ability to estimate hospital fixed effects, except of course in the two-stage

estimation procedure. While this allows us to exploit the weaker rank condition of the random effects estimator, all specifications include three variables to control for the types of hospitals in the sample. These include teaching hospitals in both Scotland and England, and Foundation Trusts (FTs) hospitals in England, which have a degree of independent self-management power compared to normal NHS hospitals, and Independent Sector Treatment Centres (ISTCs), essentially privately owned specialised treatments centres located only in England. ISTCs were introduced from 2005 onwards, and many were specifically created to provide Hip Replacement. FTs were first introduced in 2004, but each year more trusts gained this status. The dummy variable reflects this, and is "turned-on" the year a Trust earns Foundation status.

A number of further robustness checks are also run against these specifications. To ensure that the trend in activity is similar across the two countries prior to the introduction of the incentive; by running the same models on the sample from 1996-2003, the years before PbR was introduced in England, using 2002 as the falsified 'policy-on' date. We also test the basic specification excluding 'centres of excellence' (London, Glasgow and Edinburgh) in the two countries, to control for instances where uncemented activity may reflect "innovative" activity as opposed to motivation through financial incentive. Finally we also run two specifications of the standard difference-in-difference estimator, separately for the emergency and elective Hip Replacement activity to identify whether the effect is significant for each group individually, given the difference in potential cost savings between them and the presumed greater management flexibility in choice over prosthesis type when considering elective cases[11].

5. Results

Our sample on average contained 230 hospitals in any year, (194 English and 36 Scottish), that performed Hip Replacement as identified by OPCS procedure codes over the year 1996 to 2012. As not all hospitals performed Hip Replacements every year we have an unbalanced panel sample of about 4,000 observations. In reporting results we concentrate on the main coefficients of interest. Table 3 presents our OLS estimates of the difference-in-difference estimator using the main specifications outlined above. Our main coefficient of interest is the coefficient on the interaction between the PbR dummy and our country variable (coded with England=1). This is coefficient γ and is the difference-in-difference estimate of the effect of the introduction of the PbR reimbursement after 2005 for the proportion of uncemented Hip Replacement in England.

< Insert Table 3 here >

The first four columns of Table 3, present the results as estimated by equation (1). In all specifications of this model, the coefficient is appropriately signed and highly significant, indicating that after the introduction of PbR in England, as compared to Scotland, the financial incentive operated to increase the uptake of the relatively more expensive Hip prosthesis, even although NICE clinical guidance favoured the less expensive substitute. While the difference-in-difference estimate is small, with the result suggesting that each hospital in England increased its proportion of uncemented hip prostheses by 0.2 per year compared to the period prior to the introduction of PbR and in the control setting (Scotland)

[11] In both countries elective procedures dominate and are similar across the timeframe studied (approximately 89% in England and 70% in Scotland).

where the financial incentive was not introduced. The coefficient is robust to the exclusion of control variables, as also seen in Table 3.

Finally, the last two columns of Table 3, present the results from OLS estimates difference-in-trends estimator using the linear splines specification. Our main coefficients of interest are those on the interaction between the time trend variables (2005-2012 trend and 2009-2012 trend) and our country variable (coded with England=1). These are coefficients η and θ, which report the difference-in-trends estimate of the effect of the introduction of the PbR reimbursement after 2005, and its removal in 2009, for uncemented Hip Replacement in England. Coefficient η is positive, and suggests that each hospital in England increased it's up-take of uncemented hip prostheses by 5.1 per year compared to the period prior to the introduction of PbR and in the control setting (Scotland) where the financial incentive was not introduced. Coefficient θ is negative, and suggests that each hospital in England decreased it's up-take of uncemented hip prostheses by 4.1 per year compared to the period when the differential incentive was in place and in the control setting (Scotland). This implies that across the 257 English hospitals in our sample in 2009, an additional 1,311 uncemented Hip Replacements were performed as a result of the financial incentive that would not have been performed otherwise.

Table 4 presents the results as estimated by equation (4), from the two-stage model. In this specification the coefficient of interest is again appropriately signed and highly significant, indicating a step-change difference in activity of around 10, suggesting that each hospital in England increased it's up-take of uncemented hip prostheses by around 10 compared to the period prior to the introduction of the incentive and relative to Scotland. This implies that across the 257 English hospitals in our sample in 2009, an additional 2,570 uncemented Hip Replacements were performed as a result of the financial incentive, controlling for patient characteristics[12].

Tables 5, 6 and 7 present the results from the additional robustness checks. We report the trend tests in Table 5, where we run the same model on the sample of years before the PbR policy was introduced (1996-2003) with a proxy PbR introduction date (2002). The results indicate that the difference-in-trend and difference-in-difference estimates are no longer significant when we use an earlier year (2002) for our break-point. We use 2003 as the prior date here to avoid any confounding of the policy on date, as by 2004 the reimbursement tariff structure for hip replacement was in place, although it was set at a common level across both prosthesis types.

Table 6, presents the results from the tests excluding the hospitals from the 'centres of excellence', which are run to ensure that the increase in uncemented procedures is not occurring only in certain geographical regions. The basic specification is tested on samples excluding hospitals based in London, London and Glasgow and London, Glasgow and Edinburgh. The result is robust to all these specifications, and remains significant at $p<0.01$. The coefficient size slightly increases when the Scottish centres are excluded, as they are largely accounting for the uncemented activity.

<Insert Table 4 here>

[12] To test this finding we also run the model on the sample 1996-2010 – the years the financial incentive is in place - and the coefficient is stable in both size and magnitude.

<Insert Table 5 here>

<Insert Table 6 here>

Finally, Table 7 explicitly tests the basic specification using the proportions elective and emergency uncemented activity to total activity, as well as the levels of elective and emergency activity as dependent variables. The treatment effect remains significant for both elective and emergency procedures, and as expected is of greater magnitude in the elective specification which accounts for more of the Hip Replacement activity. In the levels models, we include variables for the levels of the other forms of activity (cemented, emergency and elective) to see if there is any substitution away from cemented to uncemented, and away from elective to emergency – where the financial incentive is stronger. While in both cases the coefficient is negative on cemented activity, suggesting substitution from cemented to uncemented prosthesis, it is only significant in the case of elective Hip Replacements. The coefficient on the emergency and elective coefficients are positive, suggesting that cases are increasing for both types of admission, and no substitution is occurring between them.

6. Conclusion

There is a general belief that financial incentives affect clinical and hospital behaviour. There have been a number of studies which have substantiated this belief in relating the introduction of fixed price payments (DRG payments) and competition for patients to improvements in hospital quality, both in the NHS and abroad (see for example Cooper et al, 2011; Gaynor et al, 2012; Kessler and Geppert et al, 2005). The mechanism through which this operates has been open to debate, although improvements in general hospital management are evoked to uphold these findings. Some previous literature has also found distortion in clinical practice arising from the introduction of DRGs as reimbursement is tied to length of stay (Feder et al., 1987; Newhouse & Byrne, 1988; Shen, 2003; Theurl & Winner, 2007).

This is the first study to focus attention on highly substitutable procedures that are subject to different reimbursement levels, to assess whether financial incentives affect clinical practice. It has been commonly claimed as part of the managed care literature that clinical activity will shift in response to a financial incentive, but there has little rigorous, empirical evidence to uphold the claim. We have been able to employ a difference-in-difference approach to analyse the effect within the UK NHS as, while England adopted different reimbursement levels for uncemented and cemented Hip Replacements, over the same period Scotland did not. Our results suggest that the English NHS experienced much higher, relative uptake rates of the more generously reimbursed, and presumably more profitable uncemented Hip Replacements than Scotland, once PbR had been introduced in England. This increase ensued, despite the fact that clinical guidance recommending cemented Hip Replacements had been produced by NICE, which is considered a benchmark for regulating English hospital activity. The generosity of the reimbursement, with the presumed higher mark-up given published reference costs for the procedure, coupled with a centralised procurement activity appears however to have led hospitals to pursue a management policy which is at odds with the national clinical guidance over the period of study.

This is an important finding. This conclusion is specific to this particular, perhaps unique case but the finding does support the view, and provides much needed empirical evidence, that financial incentives can trump clinical guidance. It may be that the management policies

that react to financial incentives are easier to pursue where hospitals practice centralised procurement of course. We know that financial incentives guide clinical practice in other areas (Gravelle et al. 2010). What is of interest in this case study is the financial incentive appears to lead to behaviour that contradicts national clinical guidance on hip prosthesis. It could be that clinical pressure is not brought to bear in this particular case as the different types of prosthesis appear to be highly substitutable and the clinical outcomes are similar regardless of the prosthesis used. That said on the introduction of best practice traffic within England, where no reimbursement difference is maintain across the prosthesis, data show a marked trend towards cemented prosthesis and a slowing down in uncemented activity. While other reforms are taking place around this time it is difficult to define a reform, other than the change in financial incentive and the NICE clinical guidance, that impacts differentially on use of specific prothesis type. However, we are not able to rigorously analyse the effect after the removal of the incentive, as we have only two years worth of data past 2010, but where we do control for this further change in financial incentive are results hold. The findings for our study period are however clear, on the introduction of PbR in England more generous reimbursement led to a greater volume of uncemented prosthesis use, a prosthesis which is relatively more expensive than the close substitute and which was not supported by clinical guidance.

References

Abdulkarim, A., Ellanti, P., Motterlini, N. et al, (2013) " Cemented versus uncemented fixation in total hip replacement: a systematic review and meta-analysis of randomized controlled trials", *Orthopedic Reviews* 5(1):e8.

Busse, R., Geissler, A., Quentin, W. and Wiley, M. (2011) *Diagnosis-Related Groups in Europe.* McGraw-Hill International

Charlson, M. E., P. Pompei, K. L. Ales, and C. R. MacKenzie (1987). 'A new method of classifying prognostic comorbidity in longitudinal studies: development and validation'. *Journal of Chronic Diseases* 40 (5):373–383.

Cooper , Z., Gibbons, S., Jones, S and McGuire, A. (2011) 'Does Hospital Competition Save Lives? Evidence from the English NHS patient choice reforms'. *The Economic Journal*. 121(554):F228-F260.

Cutler, DM. (1995) 'The incidence of adverse medical outcomes under prospective payment'. *Econometrica*, 63:29-55.

Dafny, LS. (2005) 'How do hospitals respond to price changes?'. *American Economic Review* 95: 1525-1547.

Davies, C. and Lorgelly, P., (2013), *Hospital procurement with concentrated sellers: A case study of hip prosthesis*, ESRC Centre for Competition Policy, University of East Anglia, CCP Working Paper 13-13.

Doiron,D., Fiebig, D., and Suziedelyte, A., (2014) 'Hips and hearts: The variation in incentive effects of insurance across hospital procedures' *Journal of Health Economics*, 37, 81-97

Farrar, S., Yi, D., Sutton, M., Chalkley, M., Sussex, J. & Scott, A. (2009). 'Has payment by results affected the way that English hospitals provide care?: Difference-in-differences analysis'. *British Medical Journal*, vol 339, b3047.

Gaynor, M, Moreno-Serra, R & Propper, (2013), 'Death by Market Power: Reform, Competition, and Patient Outcomes in the National Health Service' *American Economic Journal: Economic Policy*, 5, 134-166.,

Gaynor, M. and Town, R.,, (2012) 'Competition in Health Care Markets', in *Handbook of Health Economics*, Vol. 2, T. McGuire, M.V. Pauly, and P. Pita Barros (Eds.), Amsterdam: Elsevier North-Holland

Gaynor M, Moreno-Serra R, Propper C, (2012), 'Can competition improve outcomes in UK health care? Lessons from the past two decades', *Journal of Health Services Research & Policy*, 17, 49-5

Gilman, BH. (2000), 'Hospital response to DRG refinements: the impact of multiple reimbursement incentives on inpatient length of stay'. *Health Economics* 9(4):277-294.

Gravelle H, Santos R, Siciliani L, Goudie R., (2012), *Hospital quality competition under fixed prices*, Centre for Health Economics, University of York, *CHE Research Paper 80*

Gravelle H, Sutton M, Ma A., (2010), 'Doctor behaviour under a pay for performance contract: treating, cheating and case finding?' *Economic Journal*, 120, F129-F56.

Greer, S. (2006) 'The Politics of Health Policy Divergence', in Adams, J. and Schmueker, K. (Eds.) *Devolution in Practice 2006: Public Policy Differences within the UK*, Newcastle: Institute for Public Policy Research North.

He, D., and J. M. Mellor. 2013. "Do Changes in Hospital Outpatient Payments Affect the Setting of Care?" *Health Services Research* Vol. 48, pp.1593-1616.

Ham, C. (2004) *Health Policy in Britain: The Politics and Organization of the National Health Service*. Palgrave Macmillan, 5th Edition.

Leys, C. (1999) 'The NHS After Devolution: Editorial', *The British Medical Journal*, vol. 318, pp. 1155-1156.

Mason, A., Ward, P. and Street, A. 'England: The Healthcare Resource Group system' in Busse, R., Geissler, A., Quentin, W. and Wiley, M (2011) *Diagnosis-Related Groups in Europe*. McGraw-Hill International.

Moreno-Serra and Wagstaff (2010)

National Institute for Clinical Excellence (NICE) (2000) Guidance on the Selection of Prostheses for Primary Total Hip Replacement. NICE technology appraisal guidance 2. April 2000. http://www.nice.org.uk/guidance/ta2. (Accessed 21/7/2014)

Pollock, A. (1999) 'Devolution and Health: Challenges for Scotland and Wales', *The British Medical Journal*, vol. 318, pp. 1195-1198

Propper C, 2012, 'Competition, incentives and the English NHS', *Health Economics*, Vol:21, ISSN:1057-9230, Pages:33-40

Propper C, Sutton M, Whitnall C, et al., 2010, I'ncentives and targets in hospital care: Evidence from a natural experiment', *Journal of Public Economics*, Vol:94, ISSN:0047-2727, Pages:318-335

Scottish Parliament (2004) *National Health Service Reform (Scotland) Act 2004*, Edinburgh: Queens' Printer for Scotland.

Steele D. and Cylus J., (2014) *Health Systems In Transition: United Kingdom (Scotland)*. European Observatory on Health Systems and Policies

Street, A. and D. Dawson (2002). 'Costing hospital activity: the experience with healthcare resource groups in England'. *The European Journal of Health Economics: HEPAC: Health Economics in Prevention and Care* 3 (1), 3–9. PMID: 15609112.

White, C., and T. Yee. 2013. "When Medicare Cuts Hospital Prices, Seniors Use Less Inpatient Care." Health Affairs Vol. 32, pp:1789-95.

Yates, P., S. Serjeant, G. Rushforth, and R. Middleton (2006). The relative cost of cemented and uncemented total hip arthroplasties. The Journal of Arthroplasty 21 (1), 102–105.

Tables and Figures

Figure 1: Timeline

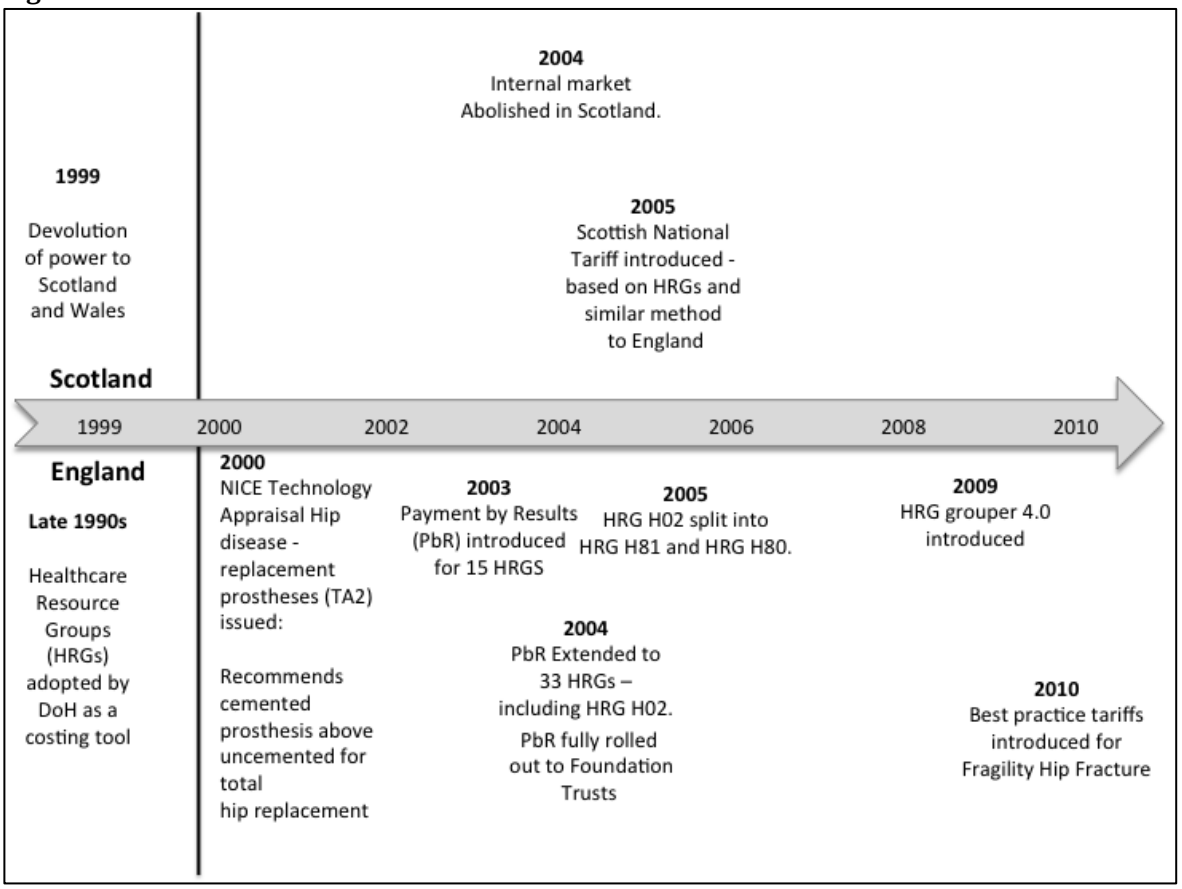


Figure 2: Cemented and Uncemented Hip Replacement cases, England and Scotland 1996-2012

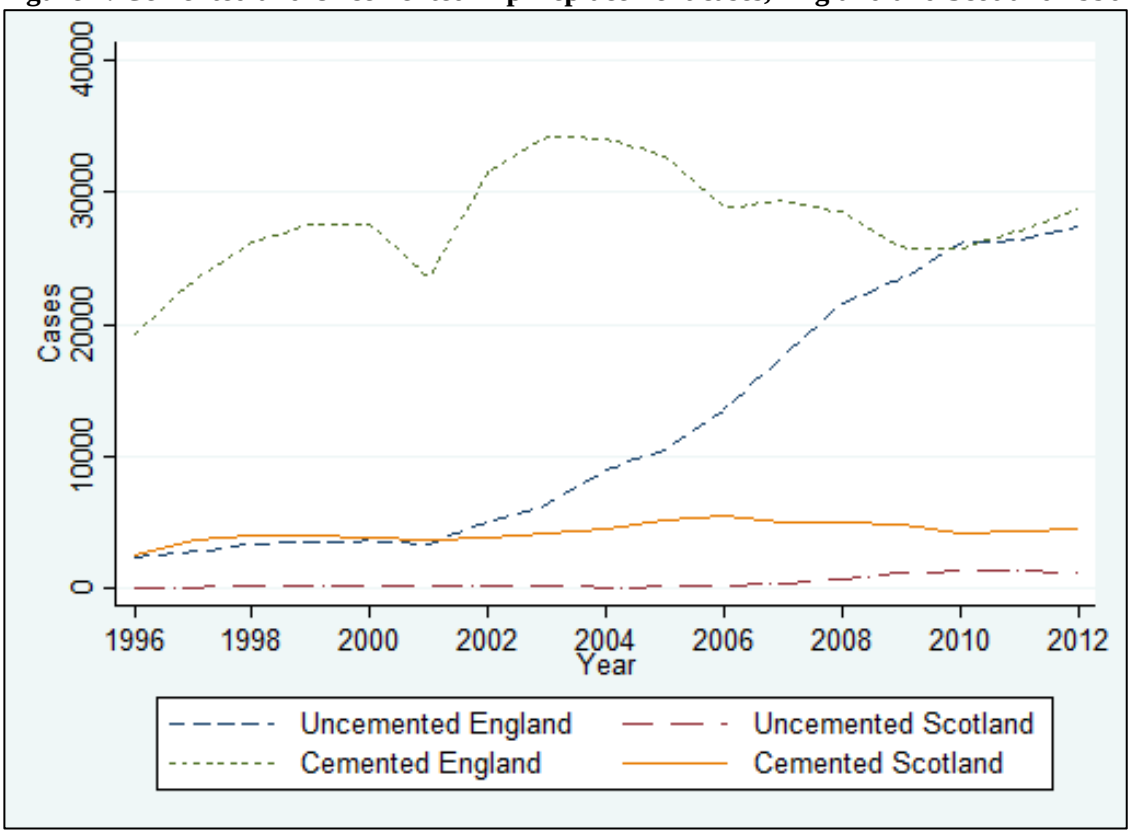
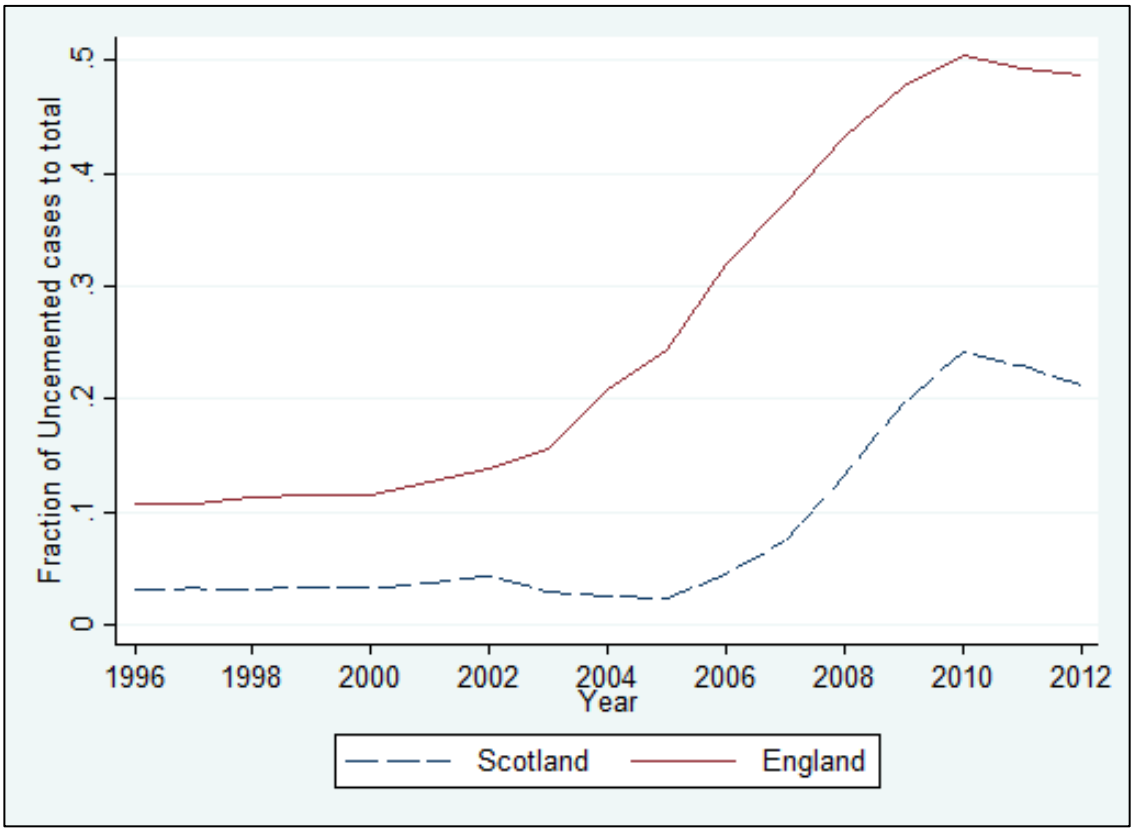


Figure 3: Ratio of Uncemented Hip Replacement cases to Total*, England and Scotland 1996-2012



*Total is measured as the sum of cemented (W371) and uncemented (W381) replacements.

Figure 4: Frequency of uncemented Hip Replacement cases to total* by provider, England and Scotland, selected years.

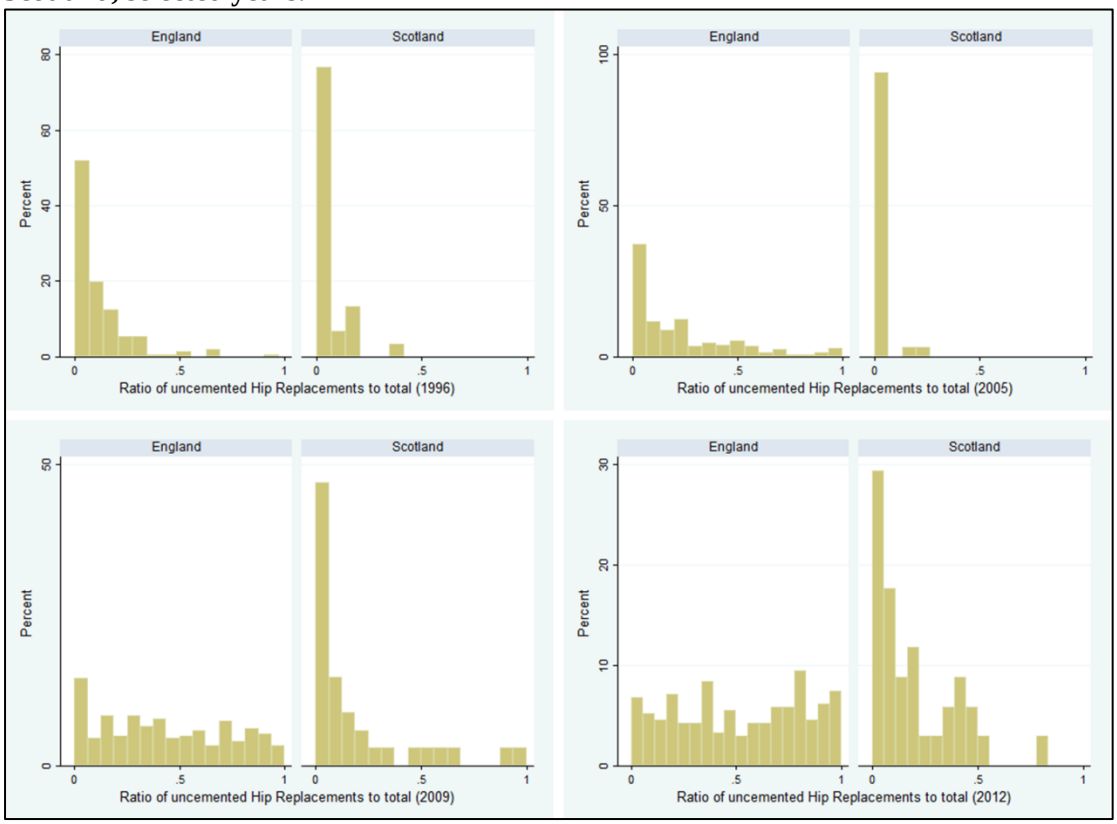


Table 1: English National Tariffs and Costs for Hip HRGs

	Year	HRG Code	Non-Elective	Elective	Estimate of Revenue gain from	
					Non-Elective Switch	Elective Switch
Spell Tariff (£)	2004/5	Primary Hip Replacement (H02)	7,776	7,776	n/a	n/a
National Average Unit Cost (£)	2001/2	Primary Hip Replacement (H02)	5,590	4,356		
Spell Tariff (£)	2005/6	Cemented (H80)	7,508	5,379	n/a	n/a
National Average Unit Cost (£)	2002/3	Primary Hip Replacement (H02)	5,010	4,660		
Spell Tariff (£)	2005/6	Uncemented (H81)	7,663	5,173		
National Average Unit Cost (£)	2002/3	Primary Hip Replacement (H02)	5,010	4,660		
Spell Tariff (£)	2006/7	Cemented (H80)	7,529	5,176	897	131
National Average Unit Cost (£)	2003/4	Cemented (H80)	4,744	4,977		
Spell Tariff (£)	2006/7	Uncemented (H81)	8,286	4,967		
National Average Unit Cost (£)	2003/4	Uncemented (H81)	4,604	4,637		
Spell Tariff (£)	2007/8	Cemented (H80)	7,717	5,305	1,482	-8
National Average Unit Cost (£)	2004/5	Cemented (H80)	5,759	5,379		
Spell Tariff (£)	2007/8	Uncemented (H81)	8,493	5,091		
National Average Unit Cost (£)	2004/5	Uncemented (H81)	5,053	5,173		
Spell Tariff (£)	2008/9	Cemented (H80)	7,304	5,220	1,692	372
National Average Unit Cost (£)	2005/6	Cemented (H80)	5,972	5,521		
Spell Tariff (£)	2008/9	Uncemented (H81)	7,816	5,587		
National Average Unit Cost (£)	2005/6	Uncemented (H81)	4,792	5,516		

Source: Department of HRG Tariff (2004/5; 2005/6; 2006/7; 2007/08; 2008/09; 2009/10); *Department of Health National Reference Costs (2002/3; 2003/4 2004/5; 2005/6; 2006/7; 2007/8; 2008/9; 2009/10)*

Table 2: Descriptive Statistics for extracted Hip Replacement Sample, Scotland and England 1996-2012

Year	Number of Hospitals		Average Age		% Male	
	England	*Scotland*	*England*	*Scotland*	*England*	*Scotland*
1996	150	34	66	71	36	30
1997	155	37	66	72	36	27
1998	155	36	66	72	37	32
1999	153	32	66	72	37	31
2000	156	34	67	72	38	32
2001	179	33	68	73	38	33
2002	190	34	68	72	38	30
2003	189	38	69	69	39	38
2004	196	36	69	70	39	33
2005	226	41	69	69	39	34
2006	180	40	69	70	37	34
2007	191	43	69	69	36	36
2008	243	24	69	70	38	36
2009	257	44	69	71	38	34
2010	291	39	69	72	38	36
2011	299	38	69	69	38	36
2012	308	40	68	71	40	39

Table 3: Results from Basic Specification

	(1) **DiD**	(2) **DiD**	(3) **DiD**	(4) **DiD**	(5) **Spline**	(6) **Spline**
Controls	Yes	No	Yes	Yes	Yes	Yes
PbR	0.239***	0.236***	0.0517	0.1652		
	(0.0400)	(0.0350)	(0.0318)	(0.0442)		
Country	0.0593**	0.0623***	0.0719**	0.0398**	0.0941***	0.0376
	(0.0253)	(0.0228)	(0.0286)	(0.0231)	(0.0244)	(0.0235)
PbR*country	0.215***	0.185***	0.155***	0.319***		
	(0.0366)	(0.0361)	(0.0362)	(0.0516)		
Year Trend					-0.00359	-0.24584
					(0.00937)	(0.01878)
2005-2009 Trend					0.0666*	0.1359*
					(0.0373)	(0.0753)
2009-2012 Trend					-0.0633	-0.1329*
					(0.0386)	(0.0712)
2005-2009 Trend*Country					0.0510***	0.0535***
					(0.0123)	(0.0202)
2009-2012 Trend*Country					-0.0406**	-0.0198**
					(0.0201)	(0.0391)
Year Dummies	Yes	Yes	No	Yes	Yes	Yes
Year Dummies*Country	No	No	No	Yes	No	Yes
Constant	0.0785**	0.0248	0.0684*	0.1016*	7.214	49.17
	(0.0367)	(0.0189)	(0.0409)	(0.0342)	(18.71)	(0.0361)
Observations	3,970	3,971	3,970	3,970	3,970	3,970
Hospitals	476	477	476	476	476	476
R-squared	0.310	0.297	0.243	0.243	0.306	0.314

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The full results for our preferred specification are reported in the Appendix. Full results for all specifications, showing coefficients on controls (which are stable across all specifications) are available on request from the authors.

Table 4: Results from Two-Stage Models

	(1) **2stage**	(2) **2stage**	(3) **2stage**	(4) **2stage**
Controls	Yes	No	Yes	Yes
PbR	0.375***	0.361***	0.206***	0.282***
	(0.0211)	(0.0199)	(0.0174)	-0.0254
Country	0.0774***	0.0620***	0.0704***	0.0644***
	(0.0173)	(0.0162)	(0.0177)	-0.0188
PbR*country	0.129***	0.106***	0.0881***	0.2406***
	(0.0230)	(0.0204)	(0.0231)	-0.0341
Year Dummies	Yes	Yes	No	Yes
Year Dummies*Country	No	No	No	Yes
Constant	0.290***	0.286***	0.282***	0.282***
	(0.0148)	(0.0135)	(0.0149)	(0.0149)
Observations	4,150	4,150	4,150	4,150
Hospitals	507	507	507	507
R-squared	0.483	0.472	0.356	0.486

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Trend Tests

	Years 1996-2003	
	DiD	**2stage**
Controls	Yes	Yes
PbR (2002)	0.0618***	-0.0301***
	(0.0173)	(0.0115)
Country	0.0589**	0.0664***
	(0.0242)	(0.0160)
Country*PbR	0.00750	-0.0174
	(0.0152)	(0.0140)
Year Dummies	Yes	Yes
Constant	0.101***	0.299***
	(0.0385)	(0.0213)
Observations	1,553	1,607
Hospitals	253	273
R-squared	0.0413	0.0651

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Controls include: age, sex, co-morbidity and hospital type for all specifications

Table 6: Geographic Test

	(1) **Excluding London**	(2) **Excluding London, Glasgow**	(3) **Excluding London, Glasgow, Edinburgh**
Controls	Yes	Yes	Yes
PbR	0.240***	0.231***	0.231***
	(0.0397)	(0.0433)	(0.0451)
Country	0.0530**	0.0498*	0.0393
	(0.0246)	(0.0267)	(0.0280)
Country*PbR	0.213***	0.226***	0.229***
	(0.0370)	(0.0396)	(0.0411)
Year Dummies	Yes	Yes	Yes
Constant	0.0865**	0.0782**	0.0778**
	(0.0378)	(0.0376)	(0.0383)
Observations	3,697	3,602	3,556
Number of h	453	446	440
R-squared	0.316	0.311	0.310

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Controls include: age, sex, co-morbidity and hospital type for all specifications

Table 7: Emergency and Elective Models

	Uncemented Proportions		*Uncemented Levels*	
	Elective	Emergency	Elective	Emergency
Cemented Levels (Elective)			-0.00457***	-0.0575
			(0.000717)	(0.0371)
Cemented Levels (Emergency)			0.0711***	-0.514
			(0.0163)	(0.397)
Unemented Levels (Elective)			0.0155***	
			(0.00341)	
Unemented Levels (Emergency)				6.583***
				(0.792)
Controls	Yes	Yes	Yes	Yes
PbR	174.7**	4.477**	0.944***	43.05***
	(78.27)	(2.254)	(0.304)	(8.345)
Country	1,052***	31.66***	0.240	14.84***
	(223.5)	(6.994)	(0.175)	(4.193)
Country*PbR	245.4**	7.399**	1.277***	51.53***
	(120.6)	(3.672)	(0.312)	(9.804)
Constant	964.1*	24.91	-0.26	-8.479
	(565.3)	(16.31)	-0.23	-5.467
Year Dummies	No	No	Yes	Yes
Observations	3,970	3,970	4,169	4,169
Hospitals	476	476	512	512
R-squared	0.201	0.199	0.373	0.347

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Controls include: age, sex, co-morbidity and hospital type for all specifications, they also include total volume for the proportions specifications.

1 **Appendix 1: Full results from Basic Specifications**

	(2) DiD	(3) DiD	(4) DiD	(5) 2stage	(6) 2stage	(7) 2stage	(6) Spline
Age 66-70	-0.0115		-0.0202				-0.00441
	(0.0190)		(0.0202)				(0.0190)
Age 71-76	-0.0566***		-0.0738***				-0.0472**
	(0.0218)		(0.0230)				(0.0217)
Age 77+	-0.0632		-0.153***				-0.0591
	(0.0571)		(0.0570)				(0.0568)
Sex	-0.0595		-0.0459				-0.0626
	(0.0543)		(0.0569)				(0.0534)
Co-morbidity	-0.00437		0.0428***				-0.00383
	(0.00915)		(0.00713)				(0.00930)
Volume	7.50e-05		0.000136*	-3.69e-05		-4.11e-05*	-0.0474*
	(7.94e-05)		(8.11e-05)	(2.50e-05)		(2.25e-05)	(0.0273)
Foundation Trust	-0.0511*		0.0107	-0.0515***		0.0452**	0.0158
	(0.0272)		(0.0266)	(0.0185)		(0.0187)	(0.0372)
Teaching Hospital	0.00338		0.00214	-0.00708		-0.00475	-0.0762**
	(0.0361)		(0.0408)	(0.0376)		(0.0349)	(0.0321)
ISTC	-0.0641**		0.0816***	-0.0653***		0.0167	
	(0.0320)		(0.0271)	(0.0237)		(0.0230)	
PbR	0.239***	0.236***	0.0517	0.375***	0.361***	0.206***	
	(0.0400)	(0.0350)	(0.0318)	(0.0211)	(0.0199)	(0.0174)	
Country	0.0593**	0.0623***	0.0719**	0.0774***	0.0620***	0.0704***	0.0985***
	(0.0253)	(0.0228)	(0.0286)	(0.0173)	(0.0162)	(0.0177)	(0.0238)
PbR*country	0.215***	0.185***	0.155***	0.129***	0.106***	0.0881***	
	(0.0366)	(0.0361)	(0.0362)	(0.0230)	(0.0204)	(0.0231)	
Year Trend							-0.00215
							(0.00936)
2005-2009 Trend							0.0638*
							(0.0374)
2009-2012 Trend							-0.0616
							(0.0386)
2005-2009 Trend*Country							0.0514***
							(0.0123)
2009-2012 Trend*Country							-0.0408**
							(0.0201)
Year Dummies	Yes	Yes	No	Yes	Yes	No	Yes
Constant	0.0785**	0.0248	0.0684*	0.290***	0.286***	0.282***	4.332
	(0.0367)	(0.0189)	(0.0409)	(0.0148)	(0.0135)	(0.0149)	(18.68)
Observations	3,970	3,971	3,970	4,150	4,150	4,150	3,970
Hospitals	476	477	476	507	507	507	476
R-squared	0.310	0.297	0.243	0.483	0.472	0.356	0.311

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2