Institute for Fiscal Studies

New joints: private providers and rising demand in the English National Health Service

IFS Working Paper W15/22

Elaine Kelly
George Stoye

New Joints: Private providers and rising demand in the English National Health Service*

Elaine Kelly and George Stoye

August 2015

Abstract

This paper investigates how changes in hospital choice sets affect levels of patient demand for elective hospital care. We exploit a set of reforms in England that opened up the market for publicly-funded patients to private hospitals. Impacts on demand are estimated using variation in distance to these private hospitals, within regions where supply constraints are fixed. We find that the reforms increased demand for publicly-funded procedures. For public hospitals, volumes remained unchanged but waiting times fell. Taken together, our results provide new insights into how individuals make choices about their care and the scope of competition between hospitals.

Keywords: health care; demand; competition
JEL Classification: I11; I14; I18; L33

*We thank the Health and Social Care Information Centre for providing access to the Hospital Episode Statistics under data sharing agreement NIC-210364-Z4K2F. This paper has been screened to ensure no confidential information is revealed. We thank the Nuffield Trust and the ESRC through The Centre for the Microeconomic Analysis of Public Policy (CPP) (ES/H021221/1) and Kelly's Future Leaders grant (ES/K009060/1) for financial support. We are grateful to James Banks, Walter Beckert, Anita Charlesworth, Richard Disney, Carol Propper, Jon Sussex, Gemma Tetlow, and participants of the Nuffield Trust Market Mechanisms in Health Seminar (2013) and the RES and EEA conferences (2014). We express particular thanks to Walter Beckert for some very useful discussions on the model. Any errors are our own. Author affiliations and contacts: Kelly (Institute for Fiscal Studies, e.kelly@ifs.org.uk); Stoye (Institute for Fiscal Studies and University College London, g.stoye@ifs.org.uk).

1 Introduction

How patients make choices about their hospital care has important implications for both hospital competition and population health. Existing work on hospital choice has focused on which hospital patients choose, taking the set of patients as given [Beckert *et al.*, 2012; Gaynor *et al.*, 2012b; Ho, 2006; Kessler and McClellan, 2000][1]. However, for most non-emergency health conditions, hospital care can be delayed or avoided. Potential patients with these conditions must therefore first decide whether to seek hospital treatment. As the relative costs and benefits of treatment will depend in part on the characteristics of hospitals available, changes in the numbers, location, or quality of local hospitals may affect not just where patients are treated, but whether patients choose to have treatment at all.

In this paper we provide evidence for how changes in the set of hospitals available to patients affects the level of demand for elective hospital care. This evidence is particularly timely, given the substantial changes in hospitals choice sets that have taken place over the past two decades, generated by hospital consolidation (see Barro and Cutler [2000]; Gaynor and Town [2012a]; Gaynor [2011] and Gaynor and Town [2012b]), the growth of managed care in the United States [Barro and Cutler, 2000; Glied, 2000; Ho, 2006][2], and reforms to increase patient choice in a number of state run European health systems[3]. The existing literatures on all these changes and reforms have typically focused upon impacts on quality, price, expenditure or productivity, rather than on demand. Yet, the response of patient demand to changes in their choice sets will affect how many patients hospitals are competing for and on what basis. Moreover, where changes in hospital choice sets affect the composition of patients, this has potentially important implications for equity and for understanding why hospital utilization varies for reasons other than medical need.

This paper considers how a set of reforms to the National Health Service (NHS) in England, which dramatically increased the number of hospitals providing publicly funded elective care, affected levels of patient demand. The majority of health care in England is funded through taxation, with hospital care typically provided through a network of state owned and run hospitals. Starting in 2003, the reforms we consider opened up the market

[1] This literature shows that demand for individual hospitals increases with quality and falls with price and distance. Distance plays a particularly important role in health care markets, as price and price differentials between hospitals are either removed in state run systems, or dampened where patients have private insurance [Kessler and McClellan, 2000].

[2] Managed care moved many insured patients from a fee-for-service insurance where there were few limits to where patients could access care, to a system where patients could choose between a network of preferred providers, thereby narrowing patient choice sets [Barro and Cutler, 2000; Glied, 2000].

[3] For example, in 2006 the English National Health Service (NHS) adopted a set of reforms that offered patients an explicit choice over where to receive hospital treatment and introduced regulated prices for hospital based on care provided [Cooper *et al.*, 2011; Gaynor *et al.*, 2013].

for NHS funded elective hospital care to privately owned hospitals [Naylor and Gregory, 2009]. By 2010/11, privately owned hospitals delivered a wide range of routine elective care for NHS-funded patients. However, in this paper we focus on the market for elective hip replacements, where the reforms acted to increase the number of hospitals where patients could receive NHS-funded procedures from 150 in 2002/03 to 257 in 2010/11[4].

Two features of the reforms are exploited to identify the impacts on patient demand. First, the reforms led to a large change in patient choice sets, increasing the number of hospitals providing elective NHS-funded hip replacements by two-thirds. The introduction these new private hospitals took place over five years and their concentration varied across the country, providing variation over both time and space. Second, we can separate the impacts on demand from those on supply by considering changes in utilization within regions where NHS administrative supply constraints remained fixed.

Our primary data source for NHS-funded procedures comes from the inpatient Hospital Episode Statistics. These administrative data document every admission to an NHS hospital, and admissions to privately-owned hospitals funded by the NHS. Each individual record contains the age, sex and local area of residence of the patient, the hospital where they were treated, details of all treatment received, and dates of admission and discharge[5].

To identify the impact of the new privately owned hospitals on demand, our empirical strategy must overcome two challenges not present when considering the choice between hospitals. The first is the absence of information on those who chose not to have a hip replacement. To address this, we aggregate numbers of NHS-funded hip replacements to the Middle Layer Super Output Area (MSOAs) level (similar to census tracts)[6]. We then construct a (balanced) panel data set of the number of procedures in all 6,781 MSOAs in England, for each financial year between April 2002 and March 2011. This panel enables us to examine changes in the levels or rates of procedure across the country, even though we cannot observe the decisions of individual patients. To exploit variation in the introduction of privately owned hospitals over time and space, we geocode each MSOA with straight-line distances to the nearest NHS and privately owned hospital in each year and augment with a series of MSOA characteristics, including demographics, other measures of population health

[4] We focus on hip replacements as (i) the procedure is performed in large volumes across most NHS hospitals in England, and (ii) privately-owned hospitals had a substantial market share of the NHS funded market by 2010/11, accounting for almost a fifth of procedures. [Kelly and Tetlow, 2012]

[5] Hip replacements are identified using the Office of Population Censuses and Surveys (OPCS) Classification of Interventions and Procedures codes (4th Edition) beginning W37, W38, W39, W93, W94 and W95. For a full list of OPCS codes see here: http://www.surginet.org.uk/informatics/opcs.php. We exclude patients admitted in an emergency, as these patients have no choice over whether they receive treatment.

[6] MSOAs are small geographic areas (with an average population of 7,200) used for census and other statistical aggregation. They have no administrative jurisdiction.

created using the Hospital Episode Statistics, and economic activity.

The second challenge to identification is separating the impacts of the new private hospitals on demand from those on supply. Here we exploit variation in the relative distance to the nearest privately owned hospital over time, within regions where supply constraints remain fixed. Existing work has demonstrated that patients show a strong preference for shorter distances [Beckert *et al.*, 2012; Gaynor *et al.*, 2012b; Kessler and McClellan, 2000], whereas administrative supply constraints over this period operated at a broader geographical level, with England divided into 152 administrative areas tasked with funding health care for their residents. To identify an impact of privately owned hospital on demand, our baseline results use the enhanced panel data set to estimate fixed effects models, which controls for time trends within the wider administrative area. Extended results then address concerns about endogenous placement of privately owned hospitals within administrative areas, by instrumenting with the presence of existing health care sites that could be used to accommodate the new hospitals. Our principal results are three-fold.

First, when privately owned hospitals become the nearest provider of NHS-funded hip replacements, MSOA annual volumes of elective hip replacements increase by 0.5 procedures per year. This compares to a pre-reform average of 5.8 hip replacements in 2002/3, and accounts for approximately 20% of the average rise of 2.6 procedures between 2002/3 and 2010/11. Using these estimates, we calculate that the additional demand for hip replacements cost the NHS £4.0 million to £9 million (6.0 - 13.5 USD) in 2010/11. Our results are robust to alternative definitions of distance and the relevant privately owned hospitals, and remain when we instrument location with pre-existing medical facilities.

Second, there is some evidence of substitution from NHS hospitals to privately owned hospitals. We estimate that the introduction of a privately owned hospital closer than the nearest NHS hospital reduces hip replacements conducted by NHS hospitals by a marginally significant 0.2 procedures per year per MSOA. At the NHS hospital level, we find that the introduction of privately owned hospitals had no effect on volumes but did act to reduce waiting times. This is consistent with a combination of some substitution from NHS to privately owned hospitals, and long waiting lists for treatment that limit the impact of such substitution on hospital volumes.

Third, over the period we consider, a negative gradient emerged in the rate of hip procedures by local area deprivation. Between 2002/03 and 2010/11, the number of hip replacements in the least deprived 10% of areas grew by 67%, almost three times the rate in the most deprived 10% of areas. While there is some evidence that privately owned hospitals contributed to the higher relative growth in the least deprived 20-40% of areas, they made almost no contribution to the growth in inequality over the remaining distribution, which was driven by NHS-funded patients treated in NHS hospitals.

Our work makes contributions to three sets of literatures. We first contribute to the literature on patient choice and modeling patient demand. In particular, we complement existing work that has estimated discrete choice models of hospital choice [Beckert *et al.*, 2012; Gaynor *et al.*, 2012b; Kessler and McClellan, 2000; Ho, 2006], by showing that for non-emergency care these types of models could be enhanced or extended by including an outside option of no procedure.[7]

Our second major contribution is to the literature on hospital competition around exit and entry. The reform we consider is in many respects the reverse of that studied in the hospital consolidation literature. Our evidence on impacts on demand therefore augments existing work on the impacts of mergers on quality, productivity, costs and prices[8]. This is particularly important as the process of health care reform and hospital consolidation remains ongoing [Dafny, 2014]. We also add to smaller literature on the entry of specialist hospitals [Barro *et al.*, 2006; Carey *et al.*, 2008; Weber, 2014], but provide evidence from a context where prices are administratively set.

Finally, we provide evidence on the more general question on how consumers make choices in markets where they first decide whether to participate and then choose which firm to pick. The most directly related markets are for other health care and social services, such as General Practitioners and vaccination clinics, and for post high school education decisions [Currie and Moretti, 2003; Manski and Wise, 1983; Mongomery, 2002].

The rest of the paper is organized as follows. Section 2 provides some institutional background and details the reforms. Section 3 presents a theoretical model and its relationship to our empirical strategy. Section 4 describes the data and our empirical method. Section 5 presents our baseline and extended results for the impacts of privately owned hospitals on demand. Section 6 considers the relationship between the ISP reforms and changes in the gradient in hip replacement provision by local area deprivation. Section 7 concludes.

[7] On a related note, we also provide evidence on the one potential mechanism which might explain the strong and pervasive relationship between hospital density and hospital utilization, known as "distance decay" [Curtis, 2004]. Existing work in the economics literature has tended to focus on how suppliers (hospitals or physicians) increase demand by exploiting their agency [Fuchs, 1978; Auster and Oaxaca, 1981; Gruber and Owings, 1996]. Here we consider how patients respond to the spatial organization of services.

[8] For example, the literature on the 2006 patient reforms in England that increased patient choice focused on clinical outcomes as measures by mortality, productivity and expenditure [Cooper *et al.*, 2011; Gaynor *et al.*, 2013]. Existing work on hospital consolidation finds that mergers dampen competition, increase prices in the free market, and have very mixed results for quality and efficiency [Town *et al.*, 2006; Dafny, 2009; Dranove and Lindrooth, 2003; Gaynor *et al.*, 2012a; Gaynor and Town, 2012b]

2 Background

2.1 Institutional Background and Policy Reforms

The majority of health care in England is funded by the state through general taxation, and provided through the National Health Service (NHS) free at the point of use. Patients access elective hospital services, such as hip replacements, through a referral from their primary care doctor or General Practitioner (GP). There are no self-referrals. Secondary or hospital care, including both outpatient consultations and inpatient treatment, has historically been delivered by state owned and run NHS Acute Trusts, or hospitals[9]. These hospitals receive per patient payments for the treatments provided, where payments for each type of treatment are set at a national level[10]. There is a small private pay sector, which accounted for a fifth of hip replacements in 2002 and are excluded from all analyses in this paper. [Arora *et al.*, 2013].

Over the past decade, two sets of NHS reforms have sought to increase patient choice and competition between hospitals. The first set of reforms gave patients a formal choice over which hospital they were referred to by their GP. From January 2006, GPs were required to offer patients a choice of four or five hospitals[11]. This replaced a system where patients could state preferences but GPs were under no obligation to actively offer their patients a choice. Henceforth, we will refer to these reforms as the "patient choice reforms". These reforms were motivated by both the belief that patients valued the choice over their care, and evidence that health care competition (when prices were fixed) could improve quality [Gaynor, 2006]. A series of work has estimated the impact of patient choice on hospital quality by comparing areas with different degrees of potential competition; this literature finds that higher degrees of competition are associated with greater improvements in quality [Cooper *et al.*, 2011; Gaynor *et al.*, 2013].

This paper focuses on a separate but related set of reforms that opened up the market for NHS-funded elective care to non-NHS providers. The NHS had purchased small volumes of care from the private or independent sector on an ad hoc basis for many years. However, reforms introduced between 2003 and 2008 formalized and greatly increased the ability of non-NHS hospitals to compete with NHS hospitals for patients, and the oppor-

[9] A NHS Acute Trust may be comprised of a single hospital or multiple hospitals within the same geographic area. For ease of expression we will refer to these NHS Acute Trusts as hospitals.

[10] Hospital care is grouped into Healthcare Resource Groups (HRGs), which are similar to Diagnostic Resource Groups in the US. Prices or Tariffs are then set at a national level based on the average cost of providing the associated care. Small adjustments are made for unavoidable local differences in costs and length of stay.

[11] The limit on the number of hospitals was then removed in 2008

tunities for NHS-funded patients to access treatment at non-NHS hospitals. We will refer to these reforms as the "Independent Sector Provider" reforms and the non-NHS hospitals as "Independent Sector Providers" (ISPs). ISPs providing hip replacements were exclusively privately owned and run hospitals.

The first wave of the ISP reforms commenced in 2003, and aimed to use ISPs to address capacity constraints within the NHS[12]. The new ISPs were to concentrate on routine patients, allowing the NHS to focus on more complex cases. Most ISPs that opened at this stage were privately owned but treated only NHS patients[13]. When the second wave was launched in 2006, the objectives expanded to include increasing competition for NHS providers and fostering innovation [Naylor and Gregory, 2009]. From this point on, most of the new ISPs were existing privately owned hospitals that treated privately and NHS funded patients alongside one another[14]. In common with the patient choice reforms, the underlying assumption was that under fixed prices increased competition would induce NHS hospitals to improve quality and efficiency. These "independent sector reforms" have received far less attention from the economics of health care literature thus far. The limited work from other disciplines has tended to focus on the type of patients treated by ISPs, rather than their impact on the market for elective care[15]. Here we consider the impacts on demand, which requires both addressing the non-random placement of ISPs with respect to potential need, and separating the effects of ISPs on demand from those on supply.

The direct effect of the ISP reforms on potential hip replacement patients was to greatly increase the set of hospitals where they could receive NHS-funded care, free at the point of use. ISPs typically offered shorter waiting times, and for some patients were closer than existing NHS hospitals. For patients treated by NHS hospitals, there were potential indirect benefits if ISPs did indeed address capacity constraints and shorten NHS hospital waiting times. For private health care companies, the independent sector reforms allowed substantial and systematic access to the NHS market for the first time in over 50 years. However, it is important to note that the payment received per patient were very similar to those received by NHS hospitals, and much lower than the price for privately funded patients. In 2010/11, the NHS payment for an elective hip replacement was approximately £5,000 (USD 7,500), or

[12] See Propper *et al.* [2010] for discussion and evidence on the waiting time targets that were implemented from 2000 onwards.

[13] These providers were known as Independent Sector Treatment Centres (ISTCs)

[14] To be clear, these hospitals were treating privately funded patients before the ISP reforms. The reforms added NHS-funded patients to their patient roll

[15] Chard *et al.* [2011]; Bardsley and Dixon [2011], for example, find that ISP patients are healthier and wealthier than those treated by NHS hospitals. However, it is important to note that this was in large part an inevitable consequence of the policy, as patients with complex medical needs could not be treated at ISPs. [Naylor and Gregory, 2009]

around half the price of a hip replacement paid for privately [Lunt *et al.*, 2011; Department of Health, 2010]. For NHS hospitals, the ISP policy at first relieved pressure to meet stringent waiting time targets, but later represented an increase in potential competition and threat to their income and market share. The extent to which ISPs retained a supporting role for some NHS hospitals with capacity constraints remains unclear.

2.2 ISPs and trends in volumes of hip replacements

Aggregate data suggest that the ISP reforms contributed to a rapid and substantive change in the level and provider composition of hip replacements. Figure 1 shows the total number of NHS-funded hip replacements conducted in each year between 2002/03 and 2010/11, by provider type. There are three points of note. First, total elective hip replacements increased by 61% over the period, only around a fifth of which can be explained by demographic changes [Arora *et al.*, 2013]. There were similar increases for most elective activity across the NHS, as funding increased and hospitals were required to meet a series of waiting time targets [Kelly and Tetlow, 2012][16]. Second, ISPs provided almost no NHS-funded hip replacements in 2002/03, but after 2007/08 their market share grew rapidly, reaching 18% in 2010/11. Third, the number of procedures conducted by NHS hospitals increased between 2002/03 and 2007/08, but leveled off after ISPs began to grow.

Figure 2 plots the locations of ISPs conducting at least 20 hip replacements per year in 2006/07, 2008/09 and 2010/11[17]. The number of ISPs grew from just nine in 2006/07, to 54 in 2008/09 and 106 in 2010/11. The map illustrates that ISPs were spread across the whole country. There was very little change in the number of NHS hospitals, which remained at around 160 throughout the period. By 2010/11, the ISP reform had therefore increased the number of hospitals available to patients by two-thirds. However, the greater volume of procedures conducted by each NHS hospital meant that NHS hospitals still delivered over 80% of all NHS-funded hip replacements by the end of the period[18].

[16]Inpatient waiting times feel from a median of 13 weeks in January 2001 to 4 weeks in January 2008, and were then constant thereafter until 2010. See Figure A.1 for details.

[17]We restrict our attention to ISPs conducting at least 20 hip replacements to avoid potential issues of disclosivity of plotting ISPs that perform more than a lower minimal threshold. Furthermore, we argue that ISPs conducting a very small number of hip replacements may not be available in the choice sets of most patients. Reducing the minimum threshold from 20 to 5 procedures increases the number of ISPs by 22%, but these smaller sites accounted for just 2.7% of ISP patients in 2010/11.

[18]In 2010/11, ISPs performed an average of 90 hip replacements per hospital, compared to over 300 per NHS hospital.

3 Understanding the relationship between ISPs and demand for NHS-funded services

In this section, we sketch a model of demand for hip replacements and use the implications of this model to develop an empirical strategy.

3.1 A theoretical model

Consider the following discrete choice problem for elective hip replacements. Patient i (in consultation with their GP) can have the procedure carried out at one of J_i hospitals, or decide not to have the procedure at all. Let U_{ij} denote patient i's indirect conditional utility from having the procedure carried out at hospital j, $j \in \mathcal{J}_i = \{1, \cdots, J_i\}$, and specify

$$U_{ij} = V_{ij} + \sigma\epsilon_{ij},$$

and,

$$V_{ij} = f(Dist_{ij}, Wait_j, Quality_j)$$

where V_{ij} denotes the measurable or quantifiable (by the econometrician) part of indirect conditional utility. Here we assume that V_{ij} is decreasing in distance from i to j ($Dist_{ij}$) and hospital waiting times ($Wait_j$), and increasing in other aspects of quality ($Quality_j$). ϵ_{ij} is an EV(0,1) residual term for each j, possibly correlated across $\mathcal{J}_i$, that captures patient-hospital level idiosyncratic tastes, and $\sigma > 0$. Here, V_{ij} and $\mathcal{J}_i$ are allowed to vary across i, through the value placed on $Dist_{ij}$. Furthermore, this specification also allows for preference heterogeneity, through patients placing different valuations on the attributes of their choice alternatives. Let $\lambda \in [0, 1]$ denote the correlation of $\epsilon_{ij} \in \mathcal{J}_i$.

Denote the indirect conditional utility of the outside option (of not undergoing the procedure) by U_{i0}, which we specify as follows

$$U_{i0} = V_0 + \sigma\epsilon_{i0},$$

and to normalize $V_0 = 0$ and assume ϵ_{i0} is also EV(0,1), independent of ϵ_{ij}, $j \in \mathcal{J}_i$. In this specification, the utility of the outside option is allowed to vary across i.[19]

This model specification implies that $\epsilon_i' = [\epsilon_{in}]_{n=0,\cdots,J}$, has a generalized extreme value distribution,

[19] This is the approach taken by Anderson and de Palma [1992]. Besanko *et al.* [1990], on the other hand, adopt a specification that restricts U_{io} to be a constant across i.

$$F(\epsilon_i) = \exp\left(-\exp(-\epsilon_{i0})\left(\sum_{j\in\mathcal{J}_i}\exp\left(-\frac{\epsilon_{ij}}{1-\lambda}\right)\right)^{1-\lambda}\right).$$

Furthermore, the probability of having the procedure carried out is governed by the inclusive value I_i of the inside options $\mathcal{J}_i$,

$$I_i = \ln\left(\sum_{j\in\mathcal{J}_i}\exp\left(\frac{V_{ij}}{\sigma(1-\lambda)}\right)\right) = \mathbb{E}\left[\max_{j\in\mathcal{J}_i} U_{ij}\right],$$

namely

$$P_1 = \Pr\left(\max_{j\in\mathcal{J}_i} U_{ij} > U_{i0}\right) = \frac{\exp\left((1-\lambda)I_i\right)}{1+\exp\left((1-\lambda)I_i\right)}.$$

Conditional on having the procedure carried out, the probability of choosing hospital j is the standard logit choice probability.

Now suppose that i's choice set is expanded, through the inclusion of ISPs in J_i. For simplicity, assume that the additional options are substitutable with regard to the present choice alternatives in $\mathcal{J}_i$.[20]. Denote i's accordingly expanded choice set by $\bar{\mathcal{J}}_i \supset \mathcal{J}_i$. Then, the inclusive value of the expanded choice set

$$\bar{I}_i = \ln\left(\sum_{j\in\bar{\mathcal{J}}_i}\exp\left(\frac{V_{ij}}{\sigma(1-\lambda)}\right)\right) = \mathbb{E}\left[\max_{j\in\bar{\mathcal{J}}_i} U_{ij}\right]$$

satisfies $\bar{I}_i > I_i$, and therefore,

$$\bar{P}_1 = \Pr\left(\max_{j\in\bar{\mathcal{J}}_i} U_{ij} > U_{i0}\right) > P_1,$$

i.e. the model predicts that, ceteris paribus, (expected) demand for the elective procedure increases, because some patients at the extensive margin will choose to have the procedure carried out in light of the expanded choice set. This is the direct effect of the policy. To the extent that capacity expansion, by inclusion of ISPs, relieves pressure on NHS hospitals in $\mathcal{J}_i$, the waiting times in the latter will decrease. If U_{ij} is decreasing in waiting times, then

[20] This can be generalized, to allow for a second, lower level nest that collects the ISPs and involves a separate correlation parameter, say $\gamma \in [0,1]$. The present model imposes the testable restriction $\gamma = \lambda$.

this indirect effect will increase V_{ij} and hence lead to a further incremental increase in I_i and $\bar{I}_i$. Who receives additional procedures will depend upon factors that vary across i and j, so in the current context where ISPs are located (through $Dist_{ij}$), and any heterogeneity in preferences for individual/hospital or hospital specific characteristics.

The intuition behind the model can be more simply explained using the supply and demand framework in Figure 3, where waiting times are used to ration procedures rather than price. The introduction of ISPs shifts the supply curve for hip replacements to the right from S_{pre} to S_{post}, resulting in a reduction in waiting times. Had this increase in supply occurred at existing hospitals, we assume that the demand curve would remain at D_{pre} and the equilibrium number of hip replacements would increase from H_0 to H_1. However, in addition to changing the supply of hip replacement beds, ISPs also change other characteristics of potential hip replacement procedures. For example, travel distances may fall or hotel amenities may be better in the ISPs. As a result, the demand curve shifts from D_{pre} to D_{post}. Unless supply is perfectly inelastic, equilibrium hip replacements will increase further from H_1 to H_2. The effect on waiting times is ambiguous and depends on the relative elasticities of demand and supply. The 'induced demand' from H_1 to H_2 differs somewhat from that estimated in the traditional supplier induced demand literature [Fuchs, 1978; Auster and Oaxaca, 1981; Gruber and Owings, 1996], where hospitals or doctors 'induce' demand, through advertising or exploiting their agency. Instead the effect captured in our model measures the impact of changing the characteristics of potential hospital treatment.

A final point to note is that the existing model does not include the option to choose privately funded treatment (approximately 20% of all elective hip replacements in 2002/03). However, the model could be adapted to allow for private pay activity, either through the outside option or formally including private hospitals in the choice set with a non-zero price. In either case there are two channels through which ISPs could increase NHS funded hip replacements: new hip replacements or substitution from the private pay to NHS-funded sector. While the overall focus of paper is on the total change in NHS-funded hip replacements, irrespective of the source, we consider potential substitution in Section 5.6.

3.2 Model to estimation

The model presented above consists of two nested choices: whether to have a hip replacement, and where to have the hip replacement. Existing work on competition and choice in the NHS has focused on the second of these decisions, and in particular on the choice over which NHS hospital to be treated by. In these papers, the econometrician observes the full sample of patients making the choice, and models are estimated using a conditional or mixed logit

[Beckert *et al.*, 2012; Gaynor *et al.*, 2012b; Ho, 2006]. The focus of this paper is on the first decision: whether to have a hip replacement at all. As such, we cannot directly estimate the nested model, as hospital administrative data only contain those who chose to receive treatment. In order to estimate whether the ISP policy affected demand on the extensive margin, we therefore adopt an alternative empirical strategy that aggregates hip replacements to the Medium Layer Super Output Area (MSOA). This allows us to consider changes in levels or rates across the country, even though we cannot observe all individuals making the choice. Existing work in this area has typically been theoretical, which may in part reflect both difficulties in obtaining a sufficiently large and measurable change in competition, and the absence of data on patients who chose the outside option[21].

A further problem in estimating the model presented is separating demand from supply. The market for elective surgery in England is subject to a range of NHS administrative constraints that do not exist in other markets. The existing model can only deal with supply constraints if they operate solely through waiting times, as longer waiting times are assumed to suppress demand. Moreover, the initial objectives of the independent sector reforms included both increasing supply and influencing demand by changing patients' choice sets.

Our estimation strategy separates demand from supply by using variation in distance to hip replacement providers within areas where supply constraints can be assumed constant. Over the period we consider, hospitals received payments for care provided to patients from Primary Care Trusts or PCTs. These organizations were in charge of funding treatment for all patients who resided within their geographic area. PCTs were established in 2002, and consolidated in number in 2006 mostly by merging old PCT areas together. From 2006 onwards, England was divided into 152 PCT areas, covering an average of 330,000 people. Of these PCT areas, 98 contained at least one MSOA where the nearest elective hip replacement provider was an ISP in 2010/11, and 107 contained at least one MSOA where the nearest elective hip replacement provider was an ISP at some point between 2002/03 and 2010/11. In other words, around two thirds of PCT areas contribute to our estimation of the impacts of the introduction of ISPs on demand.

Figure 4 illustrates our baseline approach using an example from the Leicester City PCT area. Throughout the period, elective procedures for patients living within this area were all funded by the same PCT area administration. Hospitals were paid the same tariff or price irrespective of where patients live. In 2002/03, NHS-funded hip replacements were provided by the NHS hospital Leicester General Hospital. This hospital was the nearest

[21] Brekke *et al.* [2008], for example, present a theoretical model examining the relationship between competition and waiting times in a model with a competitive segment where patients choose between hospitals, and a monopoly segment where patients choose whether to receive treatment.

NHS provider for residents within areas denoted by both the square and the circle. The ISP reforms meant that by 2010/11, patients could also receive NHS-funded treatment from Nuffield Health Leicester, which is a pre-existing privately owned hospital. For residents in the square area, the nearest provider changed to the ISP, whereas for residents of the circled area the nearest provider remained the NHS hospital. However, residents in both areas face identical administrative constraints, and differ only in the relative distance to the new provider. A larger increase in hip replacements for residents in the square area can therefore only operate through the effect of relative distance on demand.

4 Data and empirical method

This paper uses data on elective hip replacements from the inpatient Hospital Episode Statistics (HES) from April 2002 to March 2011. HES contain the records of all NHS-funded hospital care in England. This includes both care provided by NHS hospitals and care received by NHS-funded patients treated elsewhere[22]. The inpatient data contain information about the patient, including their age, sex, GP practice and local area, the admission type (emergency or elective) and dates, up to 20 diagnoses, and all procedures patients receive.[23]

To obtain a measure of how the rate of elective hip replacements varies across England and over time, we aggregate individual patient level records to the MSOA level for each financial year between April 2002 and March 2011[24]. There are a total of 6,781 MSOAs in England, with an average population of 7,200, giving a total sample of 61,029 MSOA/year observations. For each MSOA and year, we identify the nearest NHS hospital and ISP that performed at least 20 hip operations[25].

In our baseline fixed effects specification, the number of (age and sex standardized) NHS-funded hip replacements for residents of MSOA (small statistical area) m in PCT

[22] An NHS hospital (Trust) provides secondary (hospital) care to NHS patients in England. It is typically comprised of a hospital or number of hospitals run by the same hospital board, and is headquartered in the main or largest hospital.

[23] Hip replacements include those operations with Office of Population Censuses and Surveys (OPCS) Classification of Interventions and Procedures codes (4th Edition) beginning W37, W38, W39, W93, W94 and W95. Each operation code defines a different type of hip replacement. For a full list of OPCS codes see here: http://www.surginet.org.uk/informatics/opcs.php.

[24] Financial years run April to March. Emergency hip replacements are not used, because ISPs do not treat emergency patients and patients have no choice over where they are treated

[25] The nearest NHS hospital is defined by the nearest NHS Acute Trust headquarters that conducts hip replacements. The restriction to exclude ISPs conducting less than 20 procedures aims to concentrate on ISPs that are available to patients. This restriction removes 34 of the 140 ISPs that were operating in 2010/11, which accounted for just 3% of ISP patients. Robustness tests will estimate results with a minimal threshold of 5 ISPs, which account for all but 0.3% of patients.

(administrative area) p and year t is given by the following[26]:

$$Hips_{mpt} = \alpha + \beta ISP_{mt} + \gamma_m + \mu_{pt} + X_{mt} + \varepsilon_{mt} \tag{1}$$

where ISP_{mt} is a measure of the (relative) distance between the nearest ISP and the centroid of the patient's MSOA in year t. In our main specifications, ISP_{mt} is an indicator that takes the value 1 if the nearest provider of hip replacements in year t is an ISP. However, we also consider alternative specifications where we use continuous measures of the absolute distance to the nearest ISP and the distance relative to the nearest NHS hospital. The MSOA fixed effects, γ_m, control for permanent differences in the number of hip replacements across MSOAs. The PCT area specific time trends μ_{pt} control for PCT area wide factors that might change over time, and will include the average PCT area-wide effect of the introduction of an ISP. The coefficient of interest is β, the effect of introducing an ISP close to MSOA m on number of residents admitted for NHS-funded hip replacements, relative to the average effect across the patient's PCT area. This aims to capture the effect of ISPs on demand, as supply constraints are assumed to operate at the level of the PCT area, which fund the hip replacement, or the local NHS hospitals, where there may be capacity constraints.

As with any fixed effects model, the principal threat to identification is the presence of contemporaneous shocks or trends that affect the volume of hip replacements in an MSOA, which are correlated with the introduction of ISP but are not captured by PCT area wide trends. We use two principal methods to address these concerns. First, in our baseline specifications we control for MSOA level characteristics that vary over time, X_{mt}. These include: (standardized) number of admissions for fractured neck of femur and acute coronary syndrome to capture population need[27]; nearest NHS hospital emergency readmissions to hospital within 28 days of discharge to provide a measure of the quality of the nearest NHS hospital; and number of house price sales and median house price, to account for changes in economic conditions[28]. The identifying assumption is that conditional on X_{mt}

[26]NHS-funded hip replacements are age and sex standardized to reflect the England-wide population distribution in 2001

[27]Fractured neck of femur and acute coronary syndrome are emergency conditions that typically affect older people. As admissions are nearly always an emergency, admission rates should reflect patient need and be uncorrelated with the introduction of ISPs, which only treat elective cases. As with elective hip replacements, admissions are typically for older people, although the average age of admission for the two emergency conditions is somewhat older. Finally, fractured neck of femur typically results in an emergency hip replacement, which uses the same surgeons and resources as elective hip replacements. Higher rates of fractured neck of femur admissions could therefore indicate both higher need in the population, as conditions such as osteoporosis increase the need for both elective and emergency hip replacements, and greater demand on local orthopedic units from emergency patients, which could result in longer waiting times for elective patients.

[28]Population need characteristics are calculated using HES. Nearest NHS hospital characteristics are available via the NHS Health and Social Care Information Centre Indicators Portal

and μ_{pt}, other unobserved and time varying determinants of hip replacements, ε_{mt}, are uncorrelated with the introduction of ISPs. The standard error, ε_{mt}, is robust to the presence of heteroskedasticity and clustered at the PCT area level, to account for ISP placement and commissioning decisions that take place at the administrative level that funds hospital care.

Second, in Section 5.4, we address the non-random placement of ISPs using Instrumental Variables (IV) estimation, where the location of ISPs is instrumented with the presence of existing NHS and private health care facilities. It is however worth emphasizing that our baseline results are estimated *within* PCT. Any bias would therefore require that the location of ISPs within PCT areas was correlated with within PCT area variation in waiting lists or capacity constraints. Furthermore, although the introduction of the initial wave of ISPs was correlated with pent up demand, the precise timing of their introduction is likely to have been influenced by local contract or construction delays.

5 Results

5.1 Descriptive evidence on ISP use and distance

In our main specifications, where ISP_{mt} is specified as an indicator for whether the nearest provider of hip replacement is an ISP, $\widehat{\beta}$ will capture an intensity of treatment effect on demand. The 2008 patient choice reform allowed patients to choose between any provider registered to provide NHS-funded care. In that sense, all areas of England were affected or 'treated' by the introduction of ISPs. However, existing empirical evidence suggests (relative) distance is a crucial determinant of a patient's hospital choice [Beckert *et al.*, 2012; Varkevisser *et al.*, 2010; Kessler and McClellan, 2000]. Moreover, prior to the patient choice reforms, the patient's nearest NHS hospital was typically the default treatment location, accounting for two thirds of all elective hip replacements in 2002/03 [Kelly and Tetlow, 2012]. We therefore define MSOAs as 'treated' if an ISP is introduced closer than the nearest NHS hospital.

Table 1 tests whether distance is a valid measure of intensity of treatment from ISPs. Column 1 shows the proportion of MSOAs that are treated in each year between 2002/03 and 2010/11[29]. Until 2007/08, the percentage of treated MSOAs fluctuated between 2% and 4%, reflecting ad hoc purchasing from the private sector. The share of treated areas rises

(https://indicators.ic.nhs.uk/webview/). Information on house sales and prices comes from the Office for National Statistics (http://www.ons.gov.uk/ons/rel/regional-analysis/house-price-statistics-for-small-areas/1995-2013/index.html)

[29]There was no activity in 2002/03 and the small amount of activity in 2003/04 has been redacted due to the small sample size.

rapidly thereafter, reaching 22% by 2010/11. For MSOAs that were treated in 2010/11, the average distance to the nearest ISP was on average 2.4km less than the mean distance to the nearest NHS hospital (7km)[30]. In these areas, the ISP reforms therefore reduced the average distance to the nearest provider of hip replacements by a third.

For relative proximity to be a valid treatment measure, it must affect the probability that patients are treated by ISPs. The second and third columns test this assumption by comparing the probability of treatment by the nearest provider type. In all years, the share of patients treated by ISPs is higher when the nearest provider to the centriod of the MSOA is an ISP. By 2010/11, 27.0% of all NHS funded hip replacements in areas with an ISP closer were delivered by ISPs, compared to 12.5% elsewhere. This confirms that the probability that an individual is treated by an ISP declines with (relative) distance, but that NHS hospitals remain the predominant provider of NHS-funded hip replacements, even in areas with a closer ISP. The final column shows that in 2010/11 just under two-fifths of the patients treated by ISPs live nearer to an ISP than their nearest NHS hospital, illustrating that although we define treatment as living closer to an ISP than an NHS hospital, patients in other areas also receive care from ISPs.

The question addressed by this paper is not whether distance affects the probability of treatment by an ISP, but whether relative distance affects the odds that patients have surgery at all. If patients simply switched from an NHS hospital to an ISP, then there would be no increase in total volume of NHS funded hip replacements; if some patient's switching to an ISP enabled other patients to be treated by an NHS hospital, then these additional procedures should be spread evenly across their catchment area.

To assess whether the raw data provides any support for an effect of ISPs on total demand for hip replacements, Figure 5 plots the growth in the average number of NHS-funded hip replacements (the sum of those conducted by ISPs and NHS hospitals) per MSOA, by the nearest provider type in 2010/11 (NHS or ISP). Given that in 2002/03, the closest provider was always an NHS hospital, having a closer ISP in 2010/11 implies that an ISP began treating NHS-funded patients some time in the intervening period.

The solid black line shows the growth in the average number of hip replacements for the 79% of MSOAs where a NHS hospital remained the nearest provider in 2010/11. The gray line shows the growth in MSOAs where an ISP that conducted at least 20 hip replacements was the closest provider by the end of the period. There are two main points two note. First, the average number of hip replacements in the two types of area were very similar in 2002/03 at just under 6 per MSOA, before the ISP reforms were introduced, with slightly

[30] By way of comparison, the average distance between the centroid of an MSOA and the centroid of its closest MSOA is 2.3km.

higher averages in areas where a NHS hospital remained the nearest provider in 2010/11. Second, as ISP started to open, hip replacement rates grew faster in areas where an ISP was the closest provider in 2010/11: the average number of hip replacements in areas where the nearest provider was an ISP in 2010/11 rose to 9.3 (or 67%) by the end of the period, compared to 8.1 (or 40%) in areas where the nearest provider was a NHS hospital. For the dashed line, we lower the procedure threshold for an ISP from 20 to 5. As expected, the pattern is similar but somewhat weaker, as we expect that not all patients will have access to these smaller ISPs.

5.2 Fixed Effects

Table 2 provides the baseline estimates of (1). Column 1 includes only MSOA and year fixed effects, and shows that the introduction of an ISP adds 0.97 hip replacements to the MSOA average. This is statistically significant at the 1% level, and compares to an average volume of 5.8 procedures in 2002/03.

Column 2 adds PCT area specific time trends that control for all time varying factors at the PCT level, including the average impact of introducing ISPs across all MSOAs in the PCT area. This reduces the magnitude of the estimated impact of the introduction of an ISP to 0.53. The estimate remains significant at the 1% level. Adding our MSOA time-varying controls in column 3, has no effect on our coefficient of interest in terms of either magnitude or statistical significance.

An approximate cost to the NHS of the additional procedures can be calculated by combining the coefficients in Table 2 with the number of MSOAs that were treated and the NHS tariff for a hip replacement. In 2010/11, 1,471 MSOAs had an ISP conducting at least 20 hip replacements as their nearest provider. In the same year, the NHS tariff was approximately £5,000 [Department of Health, 2010]. Our most conservative estimate of the effect of ISPs on demand in column 3 therefore suggests an annual increase of 780 hips across England at a cost of £3.9 million or USD 5.9 million.

As discussed in Section 5.1, defining treatment as having an ISP as the nearest provider provides a proxy for intensity of treatment. However, intensity of treatment is not binary, and patient demand may be affected even in areas which have a nearer NHS hospital. In column 4, we divide relative distance into categories, and compare to areas where the nearest ISP is located 10km further than the nearest NHS hospital. In 2002/03 all areas fall into this category. Results indicate that the effect in areas where an ISP is the nearest provider is to increase the number of hip replacements by 0.72 hip replacements. However, there is an additional statistically significant effect for MSOAs were an ISP is less than 5 km further than the nearest NHS hospital of an extra 0.32 hip replacements. This illustrates that the effect of ISPs does decline with distance but remains statistically significant in areas where an NHS hospital is only slightly nearer. Using these coefficients, we calculate that in 2010/11 the increase in demand generated by ISPs is equal to 1,750 hip replacements at a cost of £8.75 million (USD 13.1 million).

In Appendix A, we consider different definitions of ISPs and alternative measures of treatment (including absolute and relative distance). The effect is to change the magnitude of the estimated $\widehat{\beta}$, but there remains a strong and statistically significant positive impact of ISPs on the level of demand in all specifications.

5.3 Robustness tests and extensions

The fixed effects estimates presented in Table 2 rest on two assumptions. First, that supply constraints are captured using PCT area specific time trends. Second, that the introduction of an ISP is uncorrelated with other unobserved time varying determinants of MSOA hip replacements. In Table 3, we consider alternative definitions of supply constraints. The endogeneity of ISP placement is addressed in Section 5.4.

Column 1 repeats column 3 from Table 2, where we focus on ISPs that conduct at least 20 procedures and include PCT area specific time trends. In column 2, we replace these time trends with nearest NHS hospital specific time trends. This allows for the possibility that it is the capacity at the nearest NHS hospital that dictates supply constraints in the area, as the majority of NHS hip replacement patients were still treated by their nearest hospital in 2010/11 [Kelly and Tetlow, 2012]. The estimated coefficient is smaller, but remains statistically significant at the 1% level and is not statistically different from that in column 1. Column 3 includes both PCT area and nearest NHS hospital specific time trends. Again, there is a no statistically significant change in our estimates.

The boundaries of the PCT areas did not remain constant over the period. In 2006, the number of PCT areas was reduced from 303 to 152. As a result, some MSOAs within the same PCT areas from 2006 onwards may have been subject to different capacity constraints at the beginning of the period. In column 4, we therefore apply time trends and cluster using the 2002 PCT area boundaries. The estimated $\widehat{\beta}$ falls to 0.32, but remains statistically significant at the 1% level. In Column 5, we restrict our attention to the period from April 2006 onwards, after which the boundaries of the PCT areas remained constant. Once more, the magnitude of the estimate is smaller than in column 1 but these differences are not statistically significant[31]. Taken together, this suggests that although the precise size of

[31] There is also some concern that data on ISPs is not reliable in earlier years. This specification therefore also demonstrates that any potential data issues prior to 2006/07 are not driving our baseline results.

the impact changes with the specific definition of the supply constraints used, the supply constraints that operate in a broader geographic area do not explain our ISP distance effects on the quantity demanded.

As a second check of our assumption, we consider the relationship between ISP location and the waiting times of hip replacement patients. Waiting times are the principal mechanism for allocating NHS-funded treatment. If supply constraints do operate at the broader PCT area level, there should be no within PCT area differences in waiting times for NHS or ISP services. We examine variation in waiting times for patients who had hip replacements in 2010/11, using the following specification:

$$Wait_{imp} = \theta ISPClose_{mp} + Z_{imp} + \mu_p + \epsilon_{imp} \quad (2)$$

where $Wait_{imp}$ is the waiting time of patient i in MSOA m and PCT area p who has a hip replacement in 2010/11, $ISPClose_{mp}$ is a dummy equal to one if an ISP is the closest provider in 2010/11, Z_{imp} includes the age, age squared and sex of patient i, and μ_p are PCT fixed effects. The error term ϵ_{imp} is clustered at the MSOA level and robust to the presence of heteroskedasticity. We first estimate the specification for all hip replacement patients and then split by provider type. For all hip replacement patients, we would expect the effect of having an ISP as the closest provider, $\hat{\theta}$, to be negative. As shown in Section 5.1, the probability of choosing an ISP is greater for patients that live close by, and ISPs have shorter waiting times. However, conditional on the type of provider chosen by the patient, $\hat{\theta}$ should be equal to zero.

Results in Columns 1-3 of Table 4 confirm our predictions and support the hypothesis that administrative supply constraints operate at the PCT area level. Column 1 indicates that having an ISP as your closest provider reduces waiting times by about 3 days or 0.06 standard deviations. However, for patients who choose NHS hospitals (column 2) or ISPs (column 3), there is no statistically significant difference in waiting time by relative distance to an ISP. This indicates that the result in column 1 is entirely driven by the higher probability that a patient living closer to an ISP will choose to be treated by an ISP (see Table 5.1).

The final column of Table 4, provides a preliminary test of the endogeneity of ISP placement. The specification is identical to that in column 1, except that the sample and waiting times are for hip replacement patients in 2002/03. The estimated $\hat{\theta}$ therefore indicates whether there were within PCT differences in waiting times by whether the area had an ISP as the closest provider in 2010/11. If ISPs were placed to address within PCT area capacity constraints, the estimated $\hat{\theta}$ should be positive and statistically significant. However, the results in column 4 indicates that although waiting times are higher in areas that later had an ISP, the effect is not statistically significant.

5.4 Instrumental Variable Analysis

The results in Tables 2 and 3 rest on the assumption that the introduction of an ISP is uncorrelated with other unobserved time varying determinants of MSOA hip replacements. Our estimates will be biased if, for example, ISPs were placed in areas where there were initial supply constraints or expected increases in future demand, as these areas may have seen greater rises in hip replacement numbers in the absence of the ISP reforms. This section therefore examines the determinants of ISP location in order to understand potential sources of endogeneity in our baseline estimates and assesses possible sources of random variation that could be used to 'instrument' for location.

5.4.1 The determinants of ISP placement

Table 5 shows the MSOA and nearest NHS hospital characteristics associated with having an ISP closer than the nearest NHS hospital in 2010/11. In all specifications, the dependent variable is an indicator that takes the value one if the closest provider of NHS funded hip replacements in 2010/11 was an ISP, and results are estimated using logit models. Column 1 includes measures of waiting times at the nearest NHS hospital and MSOA level, plus the average number of NHS hip replacements conducted in 2002/03 and 2004/05. These are factors that should influence location, as the initial objectives of the ISP reforms aimed to address capacity constraints and reduce waiting times. However, these determinants of location could also pose a threat to our identifying assumption, as one might expect areas with high waiting times or under-provision to experience greater growth in hip replacement volumes even in absence of the introduction of ISPs. The results indicate that a one standard deviation increase in the waiting times of the nearest hospital in 2003 increases the odds of having an ISP nearer than the nearest NHS hospital by 35%, which is in line with original stated policy objectives. However, there is no additional effect of waiting times within the MSOA, suggesting that ISP location is determined by factors at the PCT area or regional level, and not the characteristics of the much finer local area. Higher numbers of hip replacements in 2003 and 2004 are associated with slightly reduced odds of having a nearer ISP in 2010/11, which again is consistent with the objective of addressing capacity constraints.

Column 2 adds controls for socio-demographic characteristics of the MSOA and the characteristics of the nearest NHS hospital. This reduces the estimated increased odds associated with higher waiting times to 20% and slightly strengthens the association between the number of hip replacements pre-ISP and ISP location. Areas are less likely to be treated if their nearest NHS hospital was an early Foundation Trust (an indicator of higher quality) and more likely if their nearest NHS hospital was further way. The odds of having an ISP decrease with population density and the deprivation of the area.

Column 3 estimates the importance of one possible determinant of the precise location of an ISP within a PCT area: the presence of pre-existing health care facilities that could be used to accommodate ISPs. Having a private 'hospital site' nearer than the nearest NHS hospital increases the odds of having an ISP closer by thirty-fold. A NHS 'hospital site' nearer than the nearest NHS hospital conducting hip replacements doubles the odds of having an ISP, where a 'hospital site' is defined as a site with at least 30 beds and the words 'hospital' or 'infirmary' appear in the title. In most cases, this is because ISPs are located within existing NHS and private facilities. Almost all ISPs located on NHS sites are located within 'hospitals', which do not conduct hip replacements. The impact of nearest hospital waiting times is no longer statistically significant, indicating that existing health care facilities appear to be the dominant force in determining location.

The criteria for a valid instrument are (i) that the instrument must be sufficiently correlated with the endogenous variable (ISP location), and (ii) that the instrument only affects the outcome of interest (the volume of hip replacements) through its effects on ISP location. The existence of existing health care facilities certainly fulfills the first criteria, but is unlikely to fulfill the second if the outcome of interest is the volume of hip replacements in a given year, as the location of these facilities is in itself non random. However, the prospects are more promising if our outcome of interest is the change in the volume over some period of time. As the stock of existing sites is in large part fixed over the short to medium term, and their location is largely determined by historical reasons, we instead make the weaker assumption that in the absence of the ISP policy the change in NHS-funded hip replacements would be unaffected by the presence of existing private hospital sites. We therefore adopt a difference in difference style IV approach, instrumenting whether an ISP is introduced as the nearest provider with the location of pre-existing health care facilities.

5.4.2 Second stage results

To estimate the effects of the introduction of ISPs we estimate the following specification for the change in elective hip replacements between 2002/03 and 2010/11:

$$\begin{aligned}(Hip_{mp,2010}-Hip_{mp,2002}) &= \phi(ISP_{mp,2010}-ISP_{mp,2002})+(\mu_{p,2010}-\mu_{p,2002})+(\varepsilon_{mp,2010}-\varepsilon_{mp,2002}) \\ &= \phi ISP_{mp,2010} + \varphi_p + \epsilon_{mp} \quad (3)\end{aligned}$$

where our coefficient of interest is ϕ, the effect of the change in whether the nearest provider is an ISP[32], φ_p is a PCT area fixed effect, and ϵ_{mp} is the error term, which is clustered by PCT area.

Table 6 provides the estimates from our instrumental variables model. In all cases, the dependent variable is the change in the MSOA volume of NHS-funded hip replacements between 2002/03 and 2010/11. Columns 1-3 present our baseline OLS estimates, and indicate a rise of 0.82 hips per MSOA just controlling for PCT area fixed effects and 0.76 when changes in time varying controls are added to improve precision. These are larger than the estimates in Table 2 as these coefficients were an average over earlier years, when new ISPs treated fewer patients.

Column 4 presents a similar specification to column 2, but uses IV estimation. As motivated above, we use two instruments for the introduction of an ISP closer than the nearest NHS hospital by 2010/11: (i) the presence of an existing private hospital site and (ii) the presence of an existing NHS 'hospital site' that was previously unused for hip replacements. We create two indicators, which take the value of one if the respective site is closer than an NHS hospital in 2010/11, and zero otherwise. The estimated coefficient of 0.72 is slightly smaller than in column 2 but the difference is not statistically significant. Controlling for time varying MSOA characteristics in column 5 makes very little difference to our estimates. Taking the IV estimate in column 4 would imply that ISPs generated an additional 1,060 hip replacements in 2010/11 at a cost of £5.3 million. Further estimates using matching techniques provide similar results.

Results presented in Tables 2, 3 and 6 all indicate that the introduction of ISPs did increase demand for NHS-funded hip replacements. Even our most conservative estimates suggest that this increase cost the NHS more than £4 million per year by 2010/11. There are however two important points to note. First, our analysis only provides local average treatment effects estimates on demand, which operate through a specific measure of relative distance. This enables us to separate demand from supply, but means that we do not capture demand responses from those living in non-treated areas that constitute the majority of ISP patients (see Table 1). Second, the effect we estimate represents the combined effect of new procedures, and substitution from NHS hospitals. Establishing how the ISP reforms affect NHS hospitals is important for understanding the competitive effects of the reforms and the possible impacts on their patients (and therefore to the majority of patients overall). It is to this issue that we now turn.

[32] This is necessarily equivalent to whether an ISP was the closest provider in 2010, as $ISP_{mp,2002}$ is always equal to zero.

5.5 The impact of ISPs on NHS hospitals

In this section, we provide some basic evidence on the impacts of the ISP reforms on NHS hospitals on their patients. Table 7 re-estimates our baseline specification (1), but splits hip replacements into those conducted by ISPs and those conducted by NHS hospitals. The aim is to gauge the potential role of net substitution between provider types. Column 1 repeats our estimate in Column 3 of Table 2 and includes all NHS-funded hip replacements. In columns 2 and 3 we separate hip replacements conducted by ISPs from those conducted by NHS hospitals. Results indicate that it is the relative increase in hip replacements conducted by ISPs that is driving the overall effect in column 1, with the introduction of an ISP as the closest provider increasing the number of hip replacements conducted by ISPs by 0.77 procedures per MSOA per year. There is a corresponding fall in the number of hip replacements in NHS hospitals of 0.24, but the coefficient is only statistically significant at the 10% level[33]. This is consistent with some degree of net substitution between hospital types, although the negative coefficient may also be generated by higher relative rises in procedures conducted by NHS hospitals in areas where an NHS hospital remains the closest provider.

Table 7 is suggestive of some degree of substitution between NHS hospitals and ISPs, but provides no information about the overall impact on NHS hospital patient numbers or patient care. To assess the effects of ISPs on NHS hospitals and their patients, we collapse hip replacement numbers and mean waiting times of hip replacement patients to the NHS hospital year level[34]. The 9-year hospital panel contains the 133 NHS hospitals that treat more than 20 patients in all years[35]. The potential exposure of hospitals to ISPs is proxied by assigning each MSOA to their nearest NHS hospital, and calculating the share of those MSOAs in their catchment area that have a closer ISP than the NHS hospital. In 2002/03 the average share was zero, as no ISPs had opened. By 2010/11, the mean share of MSOAs within an NHS catchment area was 19.8% (median 13.7%).

Estimates presented in Table 8 consider the effect of changes in exposure to ISPs on the change in NHS hospital hip replacements and waiting times between 2002/03 and 2010/11[36]. The first set of specifications is estimated using OLS. The second set of specifications addresses the potentially endogenous placement of ISPs, by instrumenting the change in our

[33] This effect is no longer significant if we include nearest hospital fixed effects

[34] Waiting times are defined as the difference between the date of the decision to admit the patient for the procedure by the outpatient consultant and the patient being admitted to hospital for the procedure. It does not include the wait for an outpatient appointment.

[35] This compares to 159 hospitals that conducted hip replacements over the period. Most of those excluded had merged over the period we consider. All ISP patients are excluded from the analysis.

[36] As there were no ISPs in 2002/03, the change in ISP exposure between 2002/03 and 2010/11 is equal to the level of exposure in 2010/11.

exposure measure with the percentage of MSOAs within each hospital's catchment area that had a closer private hospital site in 2002/03 (see Section 5.4). The table provides two principal points of note. First, the growth of ISPs has no statistically significant impact on hip replacement volumes in NHS hospitals, either in OLS or IV. The coefficients presented in columns 1 and 2 are both imprecisely determined and close to zero[37]. The absence of an effect on NHS hospital volumes is perhaps unsurprising given that NHS treatment is rationed by waiting times rather than price and there are always patients in line. Second, the impact of growth in exposure to ISPs on waiting times is negative and statistically significant at the 5% level in both the OLS and IV estimation. The OLS estimate in column 2 indicates that moving from a zero share to the mean share of 0.2 reduced waiting times by 12 days. The IV estimates are larger, with moving to a mean share reducing waiting times by 34 days. This compares to an average reduction in waiting times of 150 days over the 8 year period we consider[38].

Taken together, results in Tables 7 and 8 are consistent with a limited degree of net substitution between NHS hospitals and ISPs. This did not reduce NHS volumes and therefore revenues, but did cut waiting times. The reduction in waiting times represents a benefit to NHS hospital patients, but also illustrates the limited potential of ISPs to exert competitive pressure on NHS hospitals when there are long lines of patients seeking treatment.

5.6 The role of the private pay sector and substitution

As noted in Section 3, the number of NHS-funded hip replacements may increase as a result of individuals switching from privately funded procedures. The private pay sector plays a relatively important role in this particular market, comprising 20% of procedures at the beginning of the period. This is one of the reasons why private capacity was first used in the orthopedic sector.

Potential substitution from the private pay to NHS-funded sector has important implications both for the interpretation and estimation of our results. In terms of interpretation, if a substantial proportion of procedures would otherwise have been conducted privately, the ISP scheme represents a transfer to patients who would otherwise have paid for the procedure. In terms of identification, the validity of our IV estimation is threatened if substitution between private and NHS-funded hip replacements was influenced by the existence of private hospital sites for reasons other than the introduction of ISPs. The recession that began in

[37]The estimated coefficients of a change of 2 hip replacements per hospital in column 1 and 4 hip replacements in column 2 compare to an average increase over the period of 83 hips replacements per hospital.

[38]This pattern of results remains unchanged if we consider the period 2006/07 to 2010/11, when we reduce the procedure threshold for ISPs to 5 procedures, and for both first and second wave ISPs.

2007 presents one such scenario. If patients in areas with existing private hospital sites were more likely to seek privately funded treatment, then we might expect the recession to increase demand for NHS treatment in those areas independently of the introduction of ISPs. However, we note that the relationship between the strength of the local housing market and hip replacements in 2, 3 and 6 is weakly positive, suggesting that this scenario is unlikely. As a further test, we restrict the sample to the period up to and including 2007, before the recession took hold. The significant effects of ISP placement remain, although the results are weaker as fewer ISPs had opened.

To gauge the potential scope of substitution we use data from the National Joint Registry (NJR), which contains information on private-pay and NHS-funded joint replacements. Figure 6 shows the number of privately financed and NHS-funded hip replacements recorded in each year between 2002/03 and 2010/11[39]. The volume of private sector hip replacements has remained roughly constant over time, with no break around the recession period. We do not claim that there has been no substitution, but that the extent is very limited[40]. This supports our claim that the location of pre independent sector reform private hospitals is a valid instrument for ISP location, as it appears unlikely that there would have been a great increase in NHS-funded procedures in areas close to private hospitals in the absence of the reform as a result of the recession.

6 The distributional impact of ISPs

Results presented in Section 5 indicate that the introduction of ISPs led to an unequal distribution of additional hip replacements across the country, with faster growth in areas close to where ISPs were located. Over the same period, Table 9 shows that a gradient emerged in the volumes of hip replacements by local area deprivation, as measured by the Index of Multiple Deprivation[41]. In 2002/03, the number of hip replacements was relatively flat across the local area deprivation distribution. Between 2002/03 and 2010/11 the average number of hip replacements increased from 5.8 to 8.4, or 46%. However, this growth was unevenly spread across the distribution. For those in the least deprived decile, average hip replacements grew by 67% almost triple the rate of hip replacements in the most deprived

[39] These data are not directly comparable to HES, as some NJR data are missing in early years (prior to 2007/8).

[40] To place this in context, the number of ISP patients in 2010/11 was greater than the sum of all privately funded patients in 2003/04

[41] The Index of Multiple Deprivation is an local area based measure of deprivation produced by the UK government that includes measures of income, employment, health deprivation and disability, education skills and training, barriers to housing and services, crime and the living environment. We use the version produced in 2004.

decile.

The set of results presented in this section seeks to understand the extent to which the introduction of ISPs can explain increases in inequality of provision. The impact of the ISP reform on equity is important for two reasons. First, equity in provision is a goal that the NHS has set for itself [National Health Service, 2015][42]. Second, waiting for a hip replacement presents costs for individuals and the rest of the health care system. For example, individuals may face additional transport costs or need more additional care; the NHS may have to pay for additional drug prescriptions or extra visits to primary care doctors. If costs incurred while waiting are positively related to need, then the way that additional procedures are allocated has implications for cost effectiveness.

6.1 ISPs and inequality in hip replacement rates by local area deprivation

As described in Section 3, ISPs may affect the level of inequality in hip replacements provision through two channels. First, differential responses to the introduction of an ISP by level of deprivation. This is plausible as patients in wealthier areas typically have fewer comorbidities and therefore have more access to ISPs. Second, an unequal distribution of where ISPs are located, and therefore which local areas are served. Results presented in Section 5.4.1 show that ISPs are typically located on private hospital sites, which tend to be located in wealthier areas.

Figure 7 considers the first channel and plots the relationship between changes in the volume of age and sex standardized hip replacements between 2002/03 and 2010/11 and MSOA deprivation. The results for all MSOAs indicate a strong negative relationship between deprivation and the change in hip replacements. Splitting by nearest provider in 2010/11 provides two points of note. First, for all levels of deprivation, the increase in the number of hip replacements is larger if the nearest provider in 2010/11 is an ISP rather than an NHS hospital. Figure 8 shows that this level effect is driven by additional operations carried out by ISPs for patients that live nearby. There are no differences in the changes in procedures that take place in NHS hospitals by nearest provider type.

Second, the deprivation gradient in the growth in hip replacements exists irrespective of whether the nearest hospital is an ISP or an NHS hospital. However, for areas where the nearest provider was an NHS hospital in 2010/11, the gradient is flat for the least deprived

[42] The NHS Constitution of 2011 states that the NHS has a "wider social duty to promote equality through the services it provides and to pay particular attention to groups or sections of society where improvements in health and life expectancy are not keeping pace with the rest of the population" [National Health Service, 2015]

40% of areas and negative thereafter. By contrast, in areas where the nearest provider is an ISP, the gradient is negative across the whole distribution. This suggests that there was some differential response to the introduction of ISPs by deprivation level, with wealthier areas benefiting more. We examine this further by repeating the analysis in Section 5 separately for each quintile of the deprivation distribution. Both the fixed effects and IV estimates indicate that the impact of introduction of ISPs on the level of demand was stronger in the least deprived fifth of areas than the most deprived fifth.[43]

To evaluate the importance of channel two, we re-weight the composition of areas so that the increases by nearest provider type are given a constant weight across the local area deprivation distribution. Figure 9 indicates that this reduces the increase for the richest 20% of areas, but makes very little difference to the rest of the distribution.

Together Figures 7 to 9 provide some evidence that ISPs contributed to the higher relative growth of hip replacements in the least deprived 20-40% of areas. However, ISPs contributed relatively little to growth inequality over the remaining distribution.

6.2 Hip replacement provision and proxies for need

The period between 2002/03 and 2010/11 was characterized by an increasing level of inequality in the distribution of hip replacements by local area deprivation. However, more relevant is the degree of imbalance or inequity in the relationship between level of provision and levels of need or health status [Wagstaff and van Doorslaer, 2000]. If the level of need varies with deprivation, then the distribution of hip replacements may need to be unequal to reflect this variation. Similarly, provision that is flat across the deprivation distribution may reflect under-provision relative to need for certain types of areas. Judge *et al.* [2010] explore equity in access to hip and knee replacement surgery by combining predicted measures of need from the English Longitudinal Study of Ageing (ELSA), a cohort study of around 10,000 people aged 50 and over in England, and measures of provision from HES. They find substantial inequity by area level deprivation, with levels of provision relative to need lower in areas with higher levels of deprivation. This suggests the fairly flat distribution of hip replacements by local area level deprivation in 2002/03 in fact represents under provision in more deprived areas. The increasing inequality seen in subsequent years implies an even greater imbalance between provision and need.

Figure 10 provides a clearer example of this by considering rates of need for joint replacements (hip and knee replacements pooled) by local area deprivation quintile using three

[43] The estimates are non-linear, with strong positive impacts also found in the middle of the deprivation distribution. However, there is also considerable noise in the estimates, with small sample sizes resulting in large standard errors.

proxy measures as reported by ELSA respondents in 2006 (ELSA wave 3): reported arthritis, whether cohort members have pain in their hips or knees, and reported difficulty in walking a quarter of a mile.[44] For all three measures there is a sharp decline in need as area level deprivation decreases. Whereas Figure 10 shows a negative relationship between wealth and need, Figure 11 indicates that individuals in the top half of the wealth distribution are more likely to report having a joint replacement in the three subsequent waves (6 years). The suggestion is therefore that the probability of having a hip replacement is not proportional to need. Changes between 2002/03 and 2010/11 indicate that there is now a reverse relationship between need and provision.

As a final check we consider how levels of admissions for fractured neck of femur have changed over the same period. These are emergency procedures to replace or repair a hip after a fracture. The average age of a fractured neck of femur patient is somewhat higher, but the same teams of surgeons will carry out both operations. In 2002/03, standardized admissions were broadly flat by local area deprivation, but highest for the most deprived areas. Between 2002/03 and 2010/11, there is a slight decrease in admissions, but no substantive changes in the distribution of admissions by local area deprivation. Conditional on immediate need for a joint replacement, there is no evidence of a gradient by local area deprivation.

7 Discussion

This paper has examined how reforms that allowed private providers access to market for NHS funded hip replacements affected the quantity of hip replacements demanded by patients. Our principal results are three fold. First, results indicate the introduction of an ISP closer than the nearest NHS hospital increases the number of hip replacements per MSOA by 0.5 hip replacements. This is equal to around 9% of the average number of 5.8 hip replacements per MSOA in pre-reform 2002/03 and 20% of average rise of 2.6 hip replacements between 2002/03 and 2010/11. Second, there is evidence of some net substitution between NHS hospital and ISPs. For NHS hospitals, the introduction of ISPs did not reduce volumes but did result in a fall in waiting times. Third, a negative gradient in hip replacement volumes by local area deprivation developed between 2002/03 and 2010/11, but this was in large part driven by patients treated by NHS hospitals rather than the introduction of ISPs. These results have the following implications.

We show that the large-scale introduction of ISPs did have an effect on demand for

[44] All three indicators strongly predict whether the cohort member has a hip replacement in the subsequent six years. We use 2006 as the base year, as this is the first wave where cohort members are asked whether their hip replacements were funded by the NHS.

elective hip replacements. For the existing hospital choice literature, this suggests suggests that discrete choice models for non-emergency procedures could be enhanced by including the outside option of no treatment. For policy makers and market regulators, our results may assist in understanding current variation in hospital utilization and inform decisions about potential changes to patient hospital choice sets. However, it is important to note that this paper does not estimate a structural model, and it is therefore not possible to conduct counter-factual policy simulations. The effects we estimate depend on the specifics of the policy reform, but also on the supply context. During the period we consider, NHS funding was increasing at unprecedented rate, allowing ISPs to gain market share without reducing volumes for NHS hospitals. Funding in the years since 2010/11 has been far more constrained, implying a supply curve with a much steeper gradient, a much smaller increase in the number of procedures performed, and greater pressures on the incomes of NHS hospitals. We are also unable to identify whether patient responses to the reform were affected by the private ownership structure of the hospitals, rather than just the availability of new sites, as there were very few changes in NHS hospitals over the period.

We add to the evidence on the nature and scope of hospital competition in three ways. First, the strength of the association between increases in demand and distance to the new ISPs re-emphasizes the importance of spatial competition in the hospital market, and provides new evidence on one mechanism behind the the strong distance decay observed in hospital utilization[45]. Second, the absence of an effect on ISPs on NHS hospital volumes shows that in systems where health care through waiting times, the impact of policies to increase competition may be limited unless there are strong constraints on supply. Third, as changes in hospital choice sets can affect demand, they also impact upon the potential revenue for competing hospitals. Changes to the organization of hospitals through consolidation need not imply that the same number of patients will reallocate to the new larger hospital.

Finally, the ISP reform changed the composition of patients receiving hip replacement treatment, as demand responses varied by distance and ISPs were not randomly allocated. Policy makers and market regulators may wish to take these potential composition changes into account when making decisions about changes to hospital organization. However, in the case of the ISP reforms, the impact on patient composition does not explain much of the negative gradient between deprivation and provision that emerged between 2002/03 and 2010/11. The source of the increasing inequality here was NHS-funded procedures conducted

[45] Cromley and McLafferty [2012] review the evidence on distance decay, and find that the effect of distance on treatment operates most strongly for elective procedures and is much weaker or absent in cases where treatment is an emergency or medically essential.

in NHS hospitals. This highlights that there are multiple mechanisms through which more economically advantaged individuals may secure better access to health care, even when health care is provided by the state. These may include access to GPs who are more willing to make referrals, or greater knowledge of the options available.

References

Anderson, S. and A. de Palma [1992]. “The Logit as a Model of product Differentiation”. *Oxford Economic Papers*, 44(51–67)

Arora, S., A. Charlesworth, E. Kelly, and G. Stoye [2013]. “Public Pay and Private Provision: The changing landscape of healthcare in the 2000s”. Nuffield Trust and Institute for Fiscal Studies Research Report

Auster, R. D. and R. L. Oaxaca [1981]. “Identification of supplier induced demand in the health care sector”. *Journal of Human Resources*, 16(3): 327–342

Bardsley, M. and J. Dixon [2011]. “Quality of care in independent sector treatment centres”. British Medical Journal 343. D6936

Barro, J. and D. M. Cutler [2000]. *Mergers and Productivity*, chap. Consolidation in the Medical Care Market Place : A Case Study from Massachusetts, pp. 9–50. University of Chicago Press

Barro, J., R. Huckman, and D. Kessler [2006]. “The effects of cardiac specialty hospitals on the cost and quality of medical care”. *Journal of Health Economics*, 25(4): 702–721

Beckert, W., M. Christensen, and K. Collyer [2012]. “Choice of NHS-Funding Hospital Services in England”. *Economic Journal*, 122(560): 400–417. DOI: 10.1111/j.1468-0297.2012.02496.x

Besanko, D., M. K. Perry, and R. H. Spady [1990]. “The Logit Model of Monopolistic Competition: Brand Diversity”. *Journal of Industrial Economics*, 38(4): 397–415

Brekke, K. R., L. Siciliani, and O. R. Straume [2008]. “Competition and waiting in hospital markets”. *Journal of Public Economics*, 97(7): 1607–1628

Carey, K., J. Burgess, and G. Young [2008]. “Specialty and full-service hospitals: A comparative cost analysis”. *Health Services Research*, 43(5): 1869–1887

Chard, J., M. Kuczawski, N. Black, and J. van der Meulen [2011]. “Outcomes of elective surgery undertaken in indepdnent sector treatment centres adn NHS providers in England: audit of patient outcomes in surgery”. British Medical Journal 343

Cooper, Z., S. Gibbons, S. Jones, and A. McGuire [2011]. "Does Hospital Competition Save Lives? Evidence From The English NHS Patient Choice Reforms". *Economic Journal*, 121(554): F228–F260. URL http://ideas.repec.org/a/ecj/econjl/v121y2011i554pf228-f260.html

Cromley, E. and S. McLafferty [2012]. *GIS and Public Health*. Guildford Press

Currie, J. and E. Moretti [2003]. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings". *The Quarterly Journal of Economics*, 118(4): 1495–1532

Curtis, S. [2004]. *Health and Inequality: Geographical Perspectives*. Sage

Dafny, L. [2009]. "Estimation and Identification of Merger Effects: An Application to Hospital Mergers". *Journal of Law and Economics*, 52(3): 523–550

Dafny, L. [2014]. "Hospital Industry Consolidation - Still More to Come?" *New England Journal of Medicine*, 370(3): 198–199. doi:10.1056/NEJMp1313948. URL http://dx.doi.org/10.1056/NEJMp1313948. PMID: 24328443

Department of Health [2010]. "NHS Payment by Results 2010-11 National Tariff Information". Department of Health

Dranove, D. and R. Lindrooth [2003]. "Hospital Consolidation and Costs: Another Look at the Evidence". *Journal of Health Economics*, 22(6): 983–997

Fuchs, V. R. [1978]. "The supply of surgeons and the demand for operations." *The Journal of Human Resources*, 13 Suppl: 35–56. doi:10.2307/145247

Gaynor, M. [2006]. "What Do We Know About Competition and Quality in Health Care Markets?" *Foundations and Trends in Microeconomics*, 2(6): 441–508. Http://www.nowpublishers.com/product.aspx?product=MIC&doi=0700000024

Gaynor, M. [2011]. "Health Care Industry Consolidation". Statement before the Committee on Ways and Means Health Subcommittee U.S. House of Representatives

Gaynor, M., M. Laudicella, and C. Propper [2012a]. "Can governments do it better? Merger mania and hospital outcomes in the English NHS". *Journal of Health Economics*, 31(3): 528–543

Gaynor, M., R. Moreno-Serra, and C. Propper [2013]. "Death by Market Power: Reform, Competition and Patient Outcomes in the National Health Service". *American Economic Journal: Economic Policy*, 5(4): 134–166. URL `http://ideas.repec.org/p/nbr/nberwo/16164.html`

Gaynor, M., C. Propper, and S. Sieler [2012b]. "Free to Choose? Reform and Demand Response in the English National Health Service." NBER Working Paper No. 18574

Gaynor, M. and R. Town [2012a]. "Competition in health care markets". In "Handbook of Health Economics", vol. 2, chap. Competition in health care markets. Elsevier

Gaynor, M. and R. Town [2012b]. "The Impact of Hospital Consolidation - Update". Robert Wood Johnson Synthesis Report, Policy Brief No. 9

Glied, S. [2000]. *Handbook of Health Economics*, vol. 1, chap. Managed Care, pp. 707–753. Elsevier

Gruber, J. and M. Owings [1996]. "Physician Financial Incentives and Cesarian Section Delivery". *RAND Journal of Economics*, 27(1): 99–123

Ho, K. [2006]. "The Welfare Effects of Restricted Hospital Choice in the US Medical Care Market". *Journal of Applied Econometrics*, 21: 1039–1079

Judge, A., N. Welton, J. Sandhu, and Y. Ben-Shlomo [2010]. "Equity in access to total joint replacement of the hip and knee in England: cross sectional study". *BMJ*, 341

Kelly, E. and G. Tetlow [2012]. "Choosing the place of care: The effect of patient choice on treatment lcoation in England, 2003-11". Nuffield Trust and Institute for Fiscal Studies Research Report

Kessler, D. P. and M. B. McClellan [2000]. "Is Hospital Competition Socially Wasteful?" *Quarterly Journal of Economics*, 115(2): 577–615

Lunt, N., R. Smith, M. Exworthy, S. T. Green, D. Horsfall, and R. Mannion [2011]. "Medical Tourism: Treatments, Markets and Health System Medical Tourism: Treatments, Markets and Health System Implications: A scoping". OECD Commissioned Report

Manski, C. and D. Wise [1983]. *College Choice in America*. Harvard University Press

Mongomery, M. [2002]. "A nested logit model of the choice of a graduate business school". *Economics of Education Review*, 21(5): 471–480

National Health Service [2015]. "Principles and values that guide the NHS". URL `http://www.nhs.uk/NHSEngland/thenhs/about/Pages/nhscoreprinciples.aspx`

Naylor, C. and S. Gregory [2009]. "Briefing: Independent Sector Treatment Centres". Kings Fund http://www.kingsfund.org.uk/publications/briefing-independent-sector-treatment-centres

Propper, C., M. Sutton, C. Whitnall, and F. Windmeijer [2010]. "Incentives and Targets in Hospital Care: Evidence from a natural experiment". *Journal of Public Economics*, 94: 318–335

Town, R., D. Wholey, R. Feldman, and L.Burns [2006]. "The Welfare Consequences of Hospital Mergers". NBER Workiing Paper No.12244

Varkevisser, M., S. van der Geest, and F. Schut [2010]. "Assessing hospital competition when prices don't matter to patients: the use of time-elasticities". *International Journal of Health Care Finance and Economics*, 10: 43–60

Wagstaff, A. and E. van Doorslaer [2000]. *Handbook of Health Economics*, chap. Equity in health care finance and delivery, pp. 1803–1862. Elsevier

Weber, E. [2014]. "Measuring Welfare from Ambulatory Surgery Centers: A Spatial Analysis of Demand for Healthcare Facilities". *Ellerie Weber*, 62(4): 591–631

A Alternative definitions of ISPs and Treatment

The first three columns of A.1 consider alternative definitions of ISPs. In the first column, we reduce the thresholds for ISPs included in our sample to five procedures. The effect is cut our coefficient of interest in half. This is unsurprising, because smaller ISPs offering very few procedures are not necessarily available to all patients. Lowering the threshold from 20 to 5 procedures increases the number of ISP sites by 22%, but these sites only treated 2.7% of all ISP patients. Columns 2 and 3 divide ISPs into Treatment Centres and Private Hospitals. Treatment Centres (or Independent Sector Treatment Centres) were established solely to treat NHS patients in the first wave of the reform. The effect of introducing a Treatment Centre is somewhat larger than our baseline estimates, at 0.67 hip replacement. "Private Hospitals", in column 3, which started operating as ISPs in the second wave of the reform were existing private hospitals that treated NHS and privately funded patients alongside one another. The estimate effect is very similar to the baseline result, in large part because most ISPs were private hospitals rather than treatment centres.

The final two columns redefine treatment as a continuous measure of distance rather than a binary indicator for relative distance. In column 4, we consider the distance from the MSOA centriod to the nearest ISP relative to the nearest NHS hospital; in column 5 we use the absolute distance to the nearest ISP. In both specifications, an additional km reduces the total number of hip replacements by 0.02.

Taken together these results indicate that altering the definition of either ISPs or MSOA treatment does effect the magnitude of our estimates. However, the main conclusion that the introduction of ISPs increases demand for elective hip replacements remains unchanged.

Table 1: ISP treatment and volumes of hip replacements 2002/03 to 2010/11

	% ISP Nearest Provider	% of MSOA hips conducted by ISPs		% ISP hips for pats with ISP closest
		NHS closest	ISP closest	
2002/03	0	0	0	0
2003/04	0	0	0	0
2004/05	2.9	1.2	4.9	10.7
2005/06	3.2	2.0	9.3	15.1
2006/07	1.9	2.9	23.6	15.1
2007/08	4.1	5.9	26.5	18.3
2008/09	12.5	7.9	22.4	30.9
2009/10	10.8	8.9	25.5	27.2
2010/11	21.7	12.5	27.0	38.7

Notes: Author's calculations using HES inpatient data April 2003 to March 2011, collapsed to the MSOA level. There was no ISP activity recorded in HES in 2002/03. Figures for 2003/04 have been omitted due to the small sample size.

Table 2: Fixed effects estimates of the effect on ISPs on standardized elective hip replacements per MSOA

	(1) MSOA FE	(2) PCT TT	(3) +controls	(4) Rel Distance
ISP location				
ISP closest provider	0.974***	0.533***	0.529***	
	(0.170)	(0.102)	(0.102)	
ISP distance relative to nearest NHS hospital				
5-10km further				0.0302
				(0.0963)
0-5km further				0.318**
				(0.132)
ISP closer				0.716***
				(0.156)
MSOA characeristics				
FNOF admits			0.0131*	0.0130*
			(0.00687)	(0.00687)
Acute coronary syndrome admits			-0.00222	-0.00232
			(0.00388)	(0.00388)
House sales			0.000780*	0.000769*
			(0.000466)	(0.000463)
Median house price (£000s)			0.00156*	0.00157**
			(0.000797)	(0.000795)
Nearest NHS Hosp characteristics				
28 day emergency readmission rate (%)			0.0142	0.0117
			(0.0357)	(0.0358)
MSOA Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
PCT area x Year Fixed Effects	No	Yes	Yes	Yes
Observations	61,029	61,029	61,029	61,029
R-squared	0.094	0.194	0.194	0.194
Number of MSOA	6,781	6,781	6,781	6,781

Notes: *** denotes significance at 1%, ** at 5%, and * at 1% level. Observations are at the MSOA year level, and the sample includes all 6781 MSOAs in England. The dependent variable in all columns is the number of admissions for an NHS-funded elective hip replacement amoungst MSOA residents, age/sex standardized to the English population in 2002. Relative distances to the nearest ISP and NHS Acute hospital are measured in straight lines from the centriod of the MSOA to the provider post (zip) code. Standard errors are robust to the presence of heteroskedasticity and clustered at the Primary Care Trust (PCT) area level.

Table 3: Robustness: Alternative definitions of supply and distance

	(1) Baseline	(2) NT TT	(3) NT&PCT TT	(4) 2002 PCT boundaries	(5) Post 2006
Sample		2002/03 - 2010/11			2006/07 - 2010/11
ISP location					
ISP closest provider	0.529***	0.477***	0.459***	0.320***	0.402***
	(0.102)	(0.106)	(0.106)	(0.0934)	(0.107)
MSOA characeristics					
FNOF admits	0.0131*	0.0161**	0.0111*	0.0110*	0.00462
	(0.00687)	(0.00641)	(0.00661)	(0.00627)	(0.0100)
Acute coronary syndrome admits	-0.00222	-0.000626	-0.00159	-0.000880	-0.00543
	(0.00388)	(0.00342)	(0.00376)	(0.00381)	(0.00484)
House sales	0.000780*	0.00137***	0.000911*	0.000652	0.00113
	(0.000466)	(0.000471)	(0.000467)	(0.000450)	(0.000682)
Median house price (£000s)	0.00156*	0.00171**	0.00151*	0.00131*	-0.00122
	(0.000797)	(0.000736)	(0.000855)	(0.000777)	(0.000976)
Nearest NHS Hosp characteristics					
28 day emergency readmission rate (%)	0.0142	-0.00657	0.0107	-0.0178	0.0705
	(0.0357)	(0.0841)	(0.0835)	(0.0403)	(0.0611)
MSOA Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
PCT area x Year Fixed Effects	Yes	No	Yes	Yes	Yes
Nrest NHS hosp X Year Fixed Effects	No	Yes	Yes	No	No
Observations	61,029	61,029	61,029	61,029	33,905
R-squared	0.194	0.192	0.229	0.232	0.129
Number of MSOA	6,781	6,781	6,781	6,781	6,781

Notes: *** denotes significance at 1%, ** at 5%, and * at 1% level. Observations are at the MSOA year level, and the sample includes all 6781 MSOAs in England. The dependent variable in all columns is the number of admissions for an NHS-funded elective hip replacement amoungst MSOA residents, age/sex standardized to the English population in 2002. In columns 1 to 4 the sample runs from financial year beginning April 2002 to financial year beginning April 2010. In column 5, the sample begins in financial year beginning April 2004 and ends in finacial year beginning in April 2010. Distances are measured in straight lines from the centriod of the MSOA to the provider post (zip) code. Standard errors are robust to the presence of heteroskedasticity and clustered at the PCT level.

Table 4: Relative distance to the nearest ISP in 2010/11 and waiting times in days

		(1)	(2)	(3)	(4)
	Year:	2010/11	2010/11	2010/11	2003/04
	Hospital type:	All	NHS	ISP	All
ISP closest hospital 2010/11		-3.503***	0.571	1.452	2.832
		(0.702)	(0.690)	(1.882)	(2.816)
Standardised effect size		-0.0603***	0.00983	0.0250	0.0166
		(0.0121)	(0.0119)	(0.0324)	(0.0165)
Observations		62,046	52,408	9,638	41,288
R-squared		0.046	0.059	0.105	0.110

Notes: *** denotes significance at 1%, ** at 5%, and * at 1% level. Observations are at the patient level. The dependent variable in all columns waiting time in days amongst patients that have hip replacements. In column 1 the sample includes all those that had an elective hip replacement between April 2010 and March 2011. Columns 2 and 3 restrict the sample to those that had a hip replacement at an NHS hospital (column 2) or an independent hospital (column 3). In column 4, the sample includes all those that had a hip replacement between April 2003 and March 2004. All specifications control a quadratic in patient age and patient sex. The standardized effect size, gives the effect size relative to overall waiting times in that year normalised to have mean zero and standard deviation one

Table 5: The odds of having an ISP nearer than the nearest NHS hospital in 2010/11

	(1)	(2)	(3)
Nrst NHS Hosp Wait 2003 (SD)	1.352***	1.195**	1.141
	(0.102)	(0.0900)	(0.117)
MSOA Wait Time 2003 (SD)	0.983	0.963	0.943
	(0.0356)	(0.0355)	(0.0451)
Average hip replacements in 2003 and 2004	0.972*	0.927***	0.939**
	(0.0159)	(0.0176)	(0.0242)
Private hospital close			29.25***
			(7.843)
NHS hospital site (>30 beds) close			2.028***
			(0.384)
Nearest NHS Hosp Characteristics			
Teaching Hosp		1.597*	0.970
		(0.419)	(0.398)
Dist (km)		1.120***	1.078***
		(0.0270)	(0.0294)
Distance sq (km)		0.997***	0.998**
		(0.000760)	(0.000768)
MSOA characteristics (2003)			
Fractured neck of femurs		0.988	0.997
		(0.0163)	(0.0226)
IMD score (2004)		0.967***	0.976
		(0.0117)	(0.0179)
Population Density		0.991***	0.990***
		(0.00284)	(0.00390)
Population		1.035	1.451
		(0.198)	(0.381)
Population aged 65-79		1.001	0.686
		(0.315)	(0.278)
Population aged 80+		2.878*	0.997
		(1.839)	(0.868)
Unemployment (2004)		1.106	1.101
		(0.0856)	(0.131)
Observations	6,710	6,710	6,710
Pseudo R-squared	0.0127	0.0731	0.404

Notes: *** denotes significance at 1%, ** at 5%, and * at 1% level. The dependent variable is an indicator equal to one if the straight line distance to the nearest ISP in 2010/11 is less than the straight line distance to the nearest hospital. Coefficients provide odds ratios from logistic specification.

Table 6: Instrumental variables estimates of the impact of ISP introduction on number of admittances for elective hip replacements per MSOA

	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV
ISP location					
Change ISP closest provider	1.392***	0.816***	0.764***	0.721***	0.626**
	(0.253)	(0.181)	(0.185)	(0.247)	(0.250)
Change in MSOA characeristics					
FNOF admits			0.0610***		0.0609***
			(0.0193)		(0.0192)
acute coronary syndrome admits			-0.0152		-0.0154
			(0.0121)		(0.0121)
House sales			0.00172		0.00173
			(0.00106)		(0.00106)
Median house price (£000s)			0.00606***		0.00614***
			(0.00177)		(0.00174)
Change in Nrest hosp characteristics					
28 day emergency readmission rate (%)			0.0540		0.0560
			(0.0437)		(0.0439)
PCT FE	No	Yes	Yes	Yes	Yes
Observations	6,781	6,781	6,781	6,781	6,781
R-squared	0.016	0.005	0.012	0.005	0.011

Notes: *** denotes significance at 1%, ** at 5%, and * at 1% level. Observations are at the MSOA year level. The dependent variable in all columns is change in the standardized number of elective hip replacements between 2002/03 and 2010/11 (standardized hip replacements in 2010/11 - standardized hip replacements in 2002/03). Relative distances to the nearest ISP and NHS hospital are measured in straight lines from the centriod of the MSOA to the provider post (zip) code. We include ISP sites that conduct at least 20 hip replacements in 2010/11. All covariates are expressed in terms of changes between 2002/03 and 2010/11. In Columns 4 and 5, whether the nearest provider was an ISP in 2010/11 is instrumented with whether there was a private or NHS hospital site nearer than the closest NHS hospital in 2002/03. Standard errors are robust to the presence of heteroskedasticity and clustered at the PCT level.

Table 7: Fixed effects estimates of the effect on ISPs on standardized elective hip replacements per MSOA, by hospital type

	(1)	(2)	(3)
Hosp Type:	All	ISP	NHS
ISP closest provider	0.529***	0.765***	-0.236*
	(0.102)	(0.116)	(0.137)
MSOA characeristics			
FNOF admits	0.0131*	0.00570*	0.00736
	(0.00687)	(0.00314)	(0.00647)
Acute coronary syndrome admits	-0.00222	0.000348	-0.00257
	(0.00388)	(0.00157)	(0.00331)
House sales	0.000780*	0.000601***	0.000180
	(0.000466)	(0.000189)	(0.000419)
Median house price (£000s)	0.00156*	6.78e-05	0.00149*
	(0.000797)	(0.000393)	(0.000772)
Nearest Trust characteristics			
28 day emergency readmission rate (%)	0.0142	0.0591**	-0.0449
	(0.0357)	(0.0264)	(0.0470)
MSOA Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
PCT area x Year Time Trend	Yes	Yes	Yes
Observations	61,029	61,029	61,029
R-squared	0.194	0.489	0.125
Number of MSOA	6,781	6,781	6,781

Notes: *** denotes significance at 1%, ** at 5%, and * at 1% level. Observations are at the MSOA year level. The sample includes all 6781 MSOAs in England, and the 9 years between April 2002 and March 2011. The dependent variable in column 1 is the number of admissions for an NHS-funded elective hip replacement amoungst MSOA residents, age/sex standardized to the English population in 2002. In columns 2 and 3, the dependent variables are the number of procedures conducted by ISPs and NHS hospitals, respectively, age and sex standardized in the same way. Standard errors are robust to the presence of heteroskedasticity and clustered at the Primary Care Trust (PCT) area level.

Table 8: Change in hip replacement patient numbers and mean waiting times in NHS hospitals and the introduction of ISPs, 2002/03 to 2010/11

	(1)	(2)	(3)	(4)
	Δ Patients		Δ Mean waiting time (days)	
	OLS	IV	OLS	IV
Δ % MSOAs ISP Closest Provider	-2.590	-4.562	-59.68**	-171.6**
	(44.13)	(106.0)	(25.65)	(66.08)
First stage F-stat	-	34	-	34
(% MSOAs Priv hosp closest 2002/03)				
Observations	133	133	133	133
R-squared	0.000	0.000	0.038	.

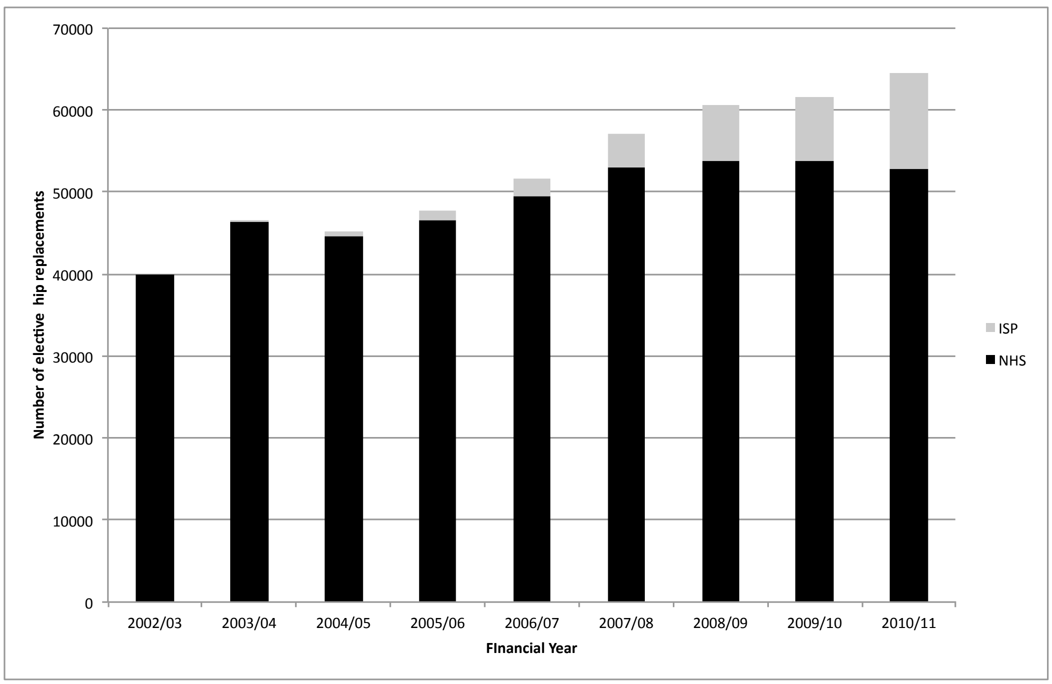
Notes: *** denotes significance at 1%, ** at 5%, and * at 1% level. Observations are at the hospital level The sample includes hospitals that data on waiting times and patient numbers in both 2002/03 and 2010/11. In columns 1 and 2, the dependent variable is the difference between the number of hip replacements conducted by the hospital in 2010/11 and the number of hip replacements conducted in 2002/03. In columns 3 and 4, the dependent variable is the mean waiting time for hip replacement patients in 2010/11 minus mean waiting times in 2002/03, where waiting times are the difference the date on which it was decided to admit the patient and the actual admission date, in days. The variable of interest in all columns is the change in % MSOAs where the ISPs is the closest provider between 2002/03 and 2010/11. This is calculated by assigning MSOAs their closest NHS hospitals, using straightline distance measures and calculating the share of these MSOAs that have a closer ISP in each year. Columns 1 and 3 are estimated using OLS. Columns 2 and 3 are estimated using IV where the change in % MSOAs where the ISPs is the closest provider is instrumented with the % of MSOA where was a private hospital site closer than the nearest NHS trust in 2002.

Table 9: Mean age/sex adjusted numbers of elective hip replacements per MSOA in 2002/03 and 2010/11, by Index of Multiple Deprivation (IMD) decile

	2002/03	2010/11	Change
All	5.8	8.4	2.6
	(3.21)	(3.8)	(4.5)
By IMD decile			
1 (Least Deprived)	5.3	8.9	3.6
2	6.0	9.5	3.5
3	6.3	9.4	3.1
4	6.1	9.3	3.1
5	6.0	9.0	3.0
6	6.4	8.7	2.3
7	5.9	8.4	2.4
8	5.5	7.4	1.9
9	5.2	7.3	2.1
10 (Most deprived)	5.0	6.2	1.1
N	6781	6781	6871

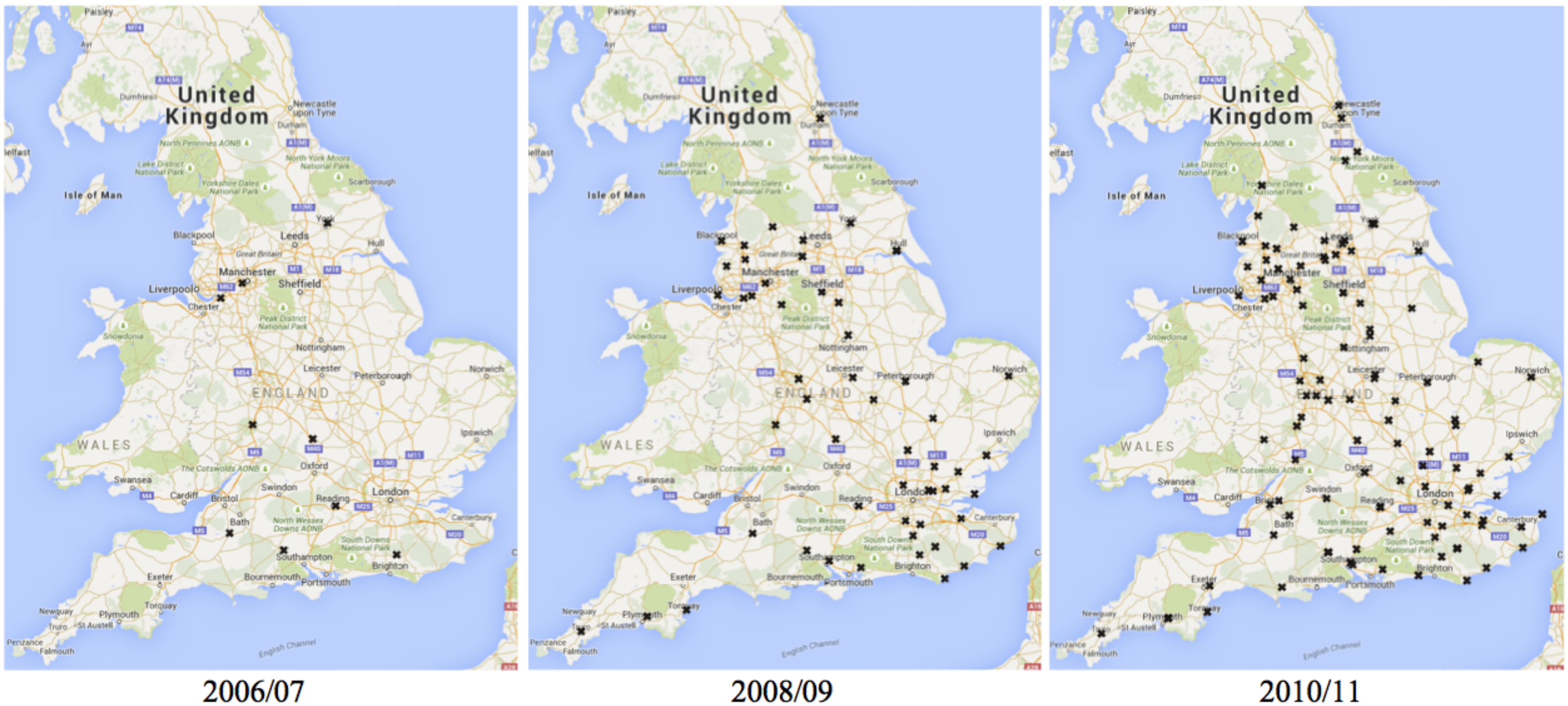
Notes: Calculated using HES inpatient data April 2002 to March 2011, collapsed to the MSOA level. Numbers of hip replacements in each MSOA are adjusted to reflect the age/sex distribution in England in 2002. Standard deviations are presented in parentheses. IMD decile is defined at the MSOA level at fixed at the 2004 value.

Figure 1: Numbers of NHS-funded hip replacements by provider type, 2002/03 to 2010/11



Notes: Hospital Episodes Statistics, April 2002- March 2011.

Figure 2: The spread of ISPs across England (2006/07, 2008/09 and 2010/11)



Notes: Sample is restricted to those Independent Sector Providers that are recorded as conducting at least 20 NHS-funded elective hip replacements in the NHS Hospital Episode Statistics in the given year

Figure 3: Independent Sector Providers and the Demand for and Supply of Hip Replacement Procedures

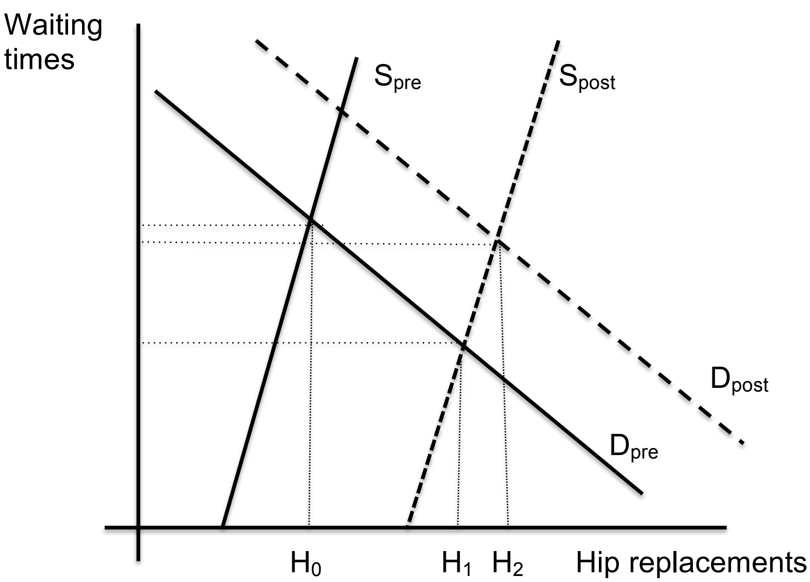


Figure 4: Separating demand from supply: An example from Leicester City Primary Care Trust

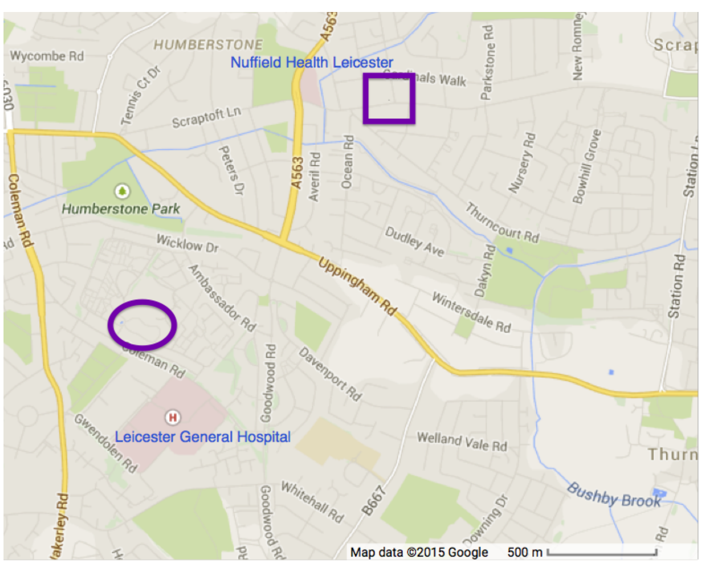
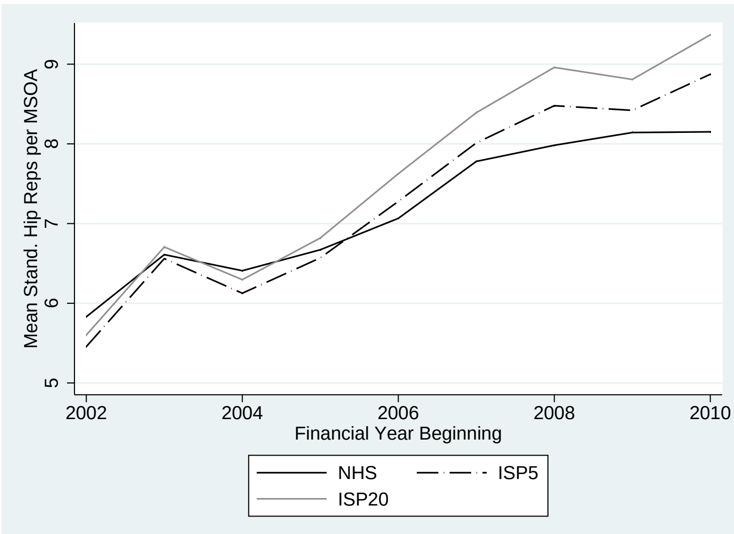
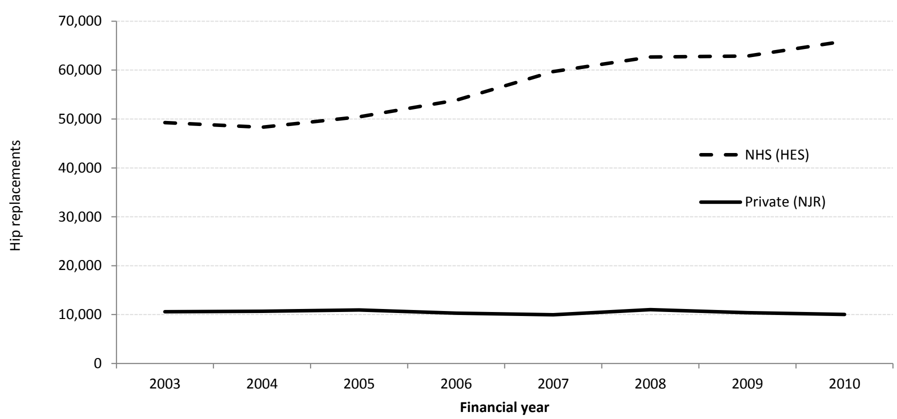


Figure 5: Mean standardized elective hip replacements per MSOA, by nearest provider type in 2010/11



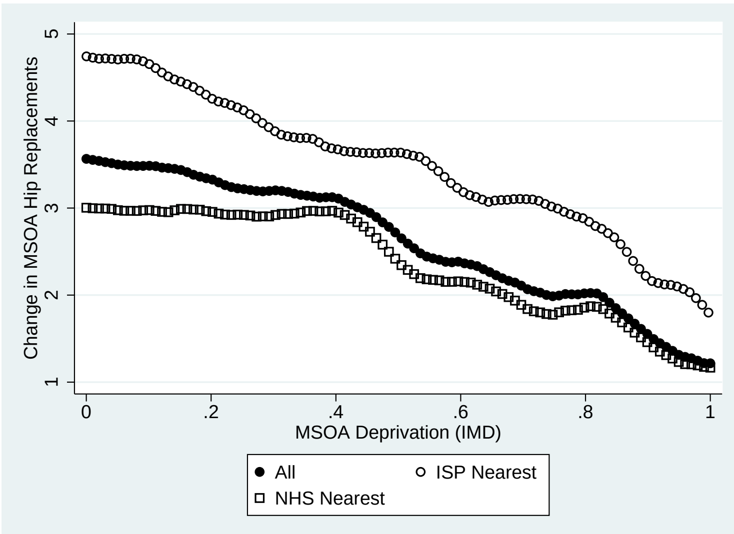
Notes: The nearest provider (NHS or ISP) is defined as the nearest hospital that conducted at least 5 (or 20) hip replacements in 2010/11. All distances are measure in a straight line from the centroid of the MSOA to the full postcode of the NHS hospital or the ISP site.

Figure 6: The volume of hip replacements recorded in the National Joint Registry and Hospital Episodes Statistics, by funding type (2003/04 to 2010/11)



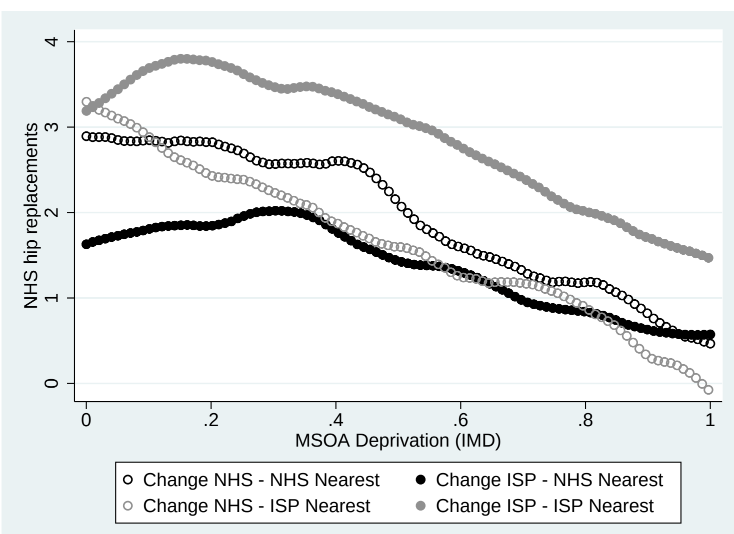
Notes and sources: Authors' calculations using HES and NJR data. HES series includes all NHS-funded procedures, conducted either by an NHS hospital or an ISP. NJR series includes all privately financed procedures, conducted in private hospitals or NHS private wings.

Figure 7: Average growth in hip replacements per MSOA between 2002/03 to 2010/11, by local area deprivation



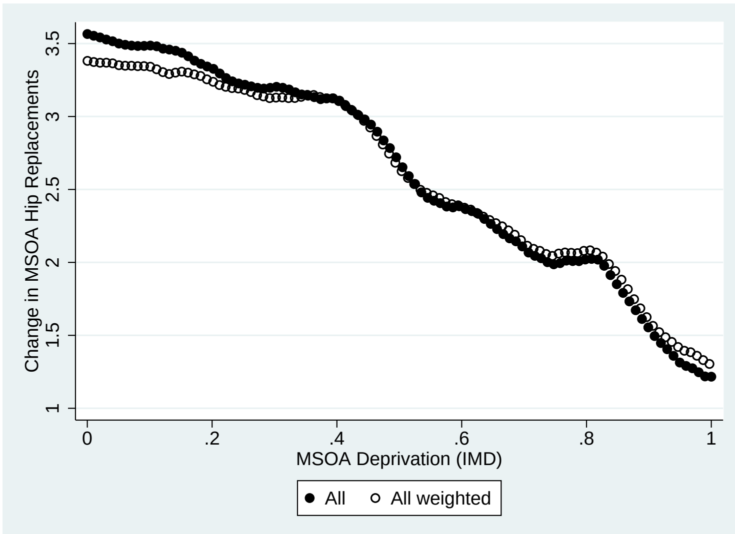
Notes: Univariate Nadaraya-Watson kernel regression of change in annual age/sex adjusted hip replacements between 2002/03 and 2010/11 by MSOA Index of Multiple Deprivation (IMD), using an epanechnikov kernel. ISPs include those sites that perform at least 20 procedures in 2010/11.

Figure 8: Change in average age/sex adjusted elective hip replacements between 2002/03 and 2010/11 by provider type, and the nearest provider type



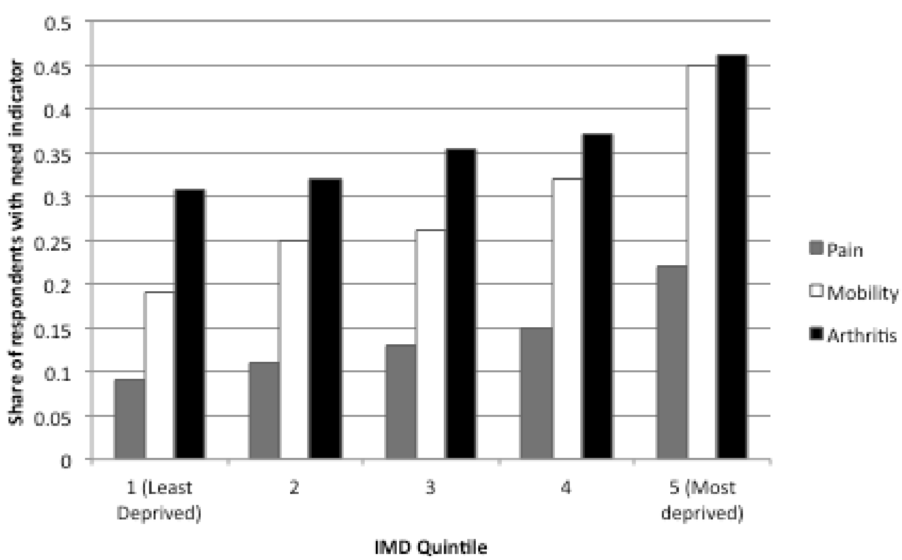
Notes: Univariate Nadaraya-Watson kernel regression of change in annual age/sex adjusted hip replacements between 2002/03 and 2010/11 by MSOA Index of Multiple Deprivation (IMD), using an epanechnikov kernel. ISPs include those sites that perform at least 20 procedures in 2010/11.

Figure 9: Weighted growth in NHS-funded hip replacements between 2002/03 and 2010/11, by local area deprivation



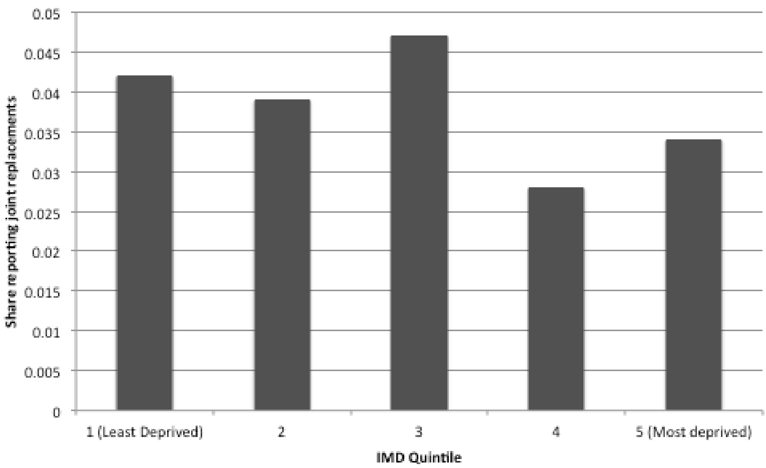
Notes: Univariate Nadaraya-Watson kernel regression of change in annual age/sex adjusted hip replacements between 2002/03 and 2010/11 by MSOA Index of Multiple Deprivation (IMD), using an epanechnikov kernel. ISPs include those sites that perform at least 20 procedures in 2010/11. ISPs are given a constant weight of 0.217 for all areas.

Figure 10: Share of respondents reporting proxies for need for a joint replacement, by wealth quintile in 2006



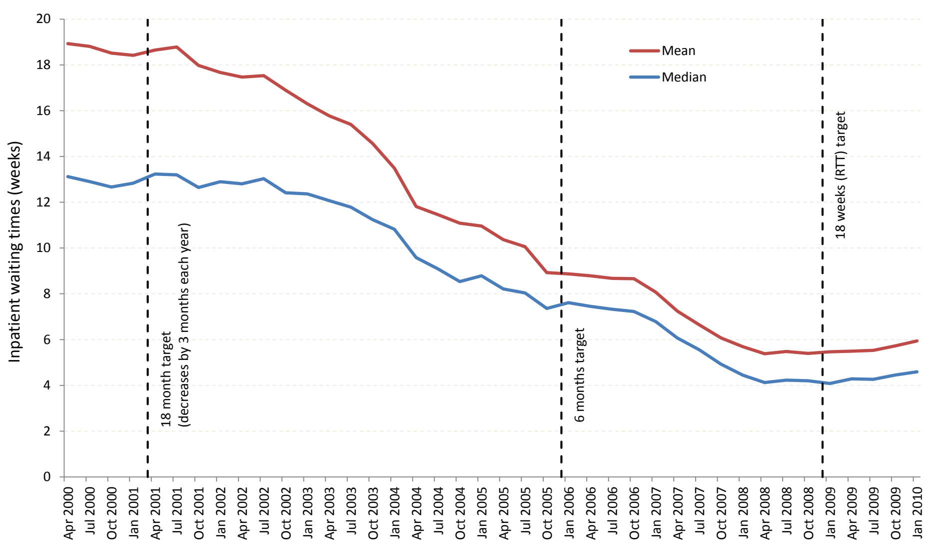
Notes: Sample include ELSA respondents in wave 3 (2006). Cohort members are categorized as having arthritis if they report ever being diagnosed with arthritis. Poor mobility denotes difficulty in walking a quarter of a mile. Cohort members are defined as having joint pain if they report pain in their hips or knees.

Figure 11: Share of ELSA respondents that reported a joint replacement in 2008, 2010, or 2012 by wealth quintile in 2006.



Notes: The sample includes ELSA respondents in wave 3 (2006). Prevalence is based on whether the cohort member reports at least one hip or knee replacement in wave 4 (2008) wave 5 (2010) or wave 6 (2012).

Figure A.1: Mean and Median Inpatient Waiting Times (April 2000-January 2010)



Source and notes: Department of Health provider-based median inpatient waiting times. Waiting times measure time elapsed between a consultant's decision to admit and the date of admission for an inpatient procedure, and does not include the time between GP referral and outpatient consultation(s). The first waiting times target was introduced in April 2001, with a maximum wait of 18 months between the decision to admit and inpatient admission. The target was reduced by three months each year. In December 2008 a new referral to treatment (RTT) target was introduced, with a maximum wait of 18 weeks between GP referral and inpatient admission.

Table A.1: Fixed effects estimates of the effect on ISPs on standardized elective hip replacements per MSOA for alternative treatment definitions

	(1) ISP5	(2) Treatment Centre	(3) Private Hosp	(4) Rel Distance	(5) Ab Distance
Sample	2002/03 - 2010/11			2004/05 - 2010/11	
ISP location					
ISP5 closest provider	0.246***				
	(0.0852)				
ISTC20 closest provider		0.667**			
		(0.268)			
AQP20 closest provider			0.499***		
			(0.114)		
Rel dist to nearest ISP (km)				-0.0161***	
				(0.00469)	
Dist to nearest ISP (km)					-0.0161***
					(0.00559)
MSOA characteristics	Yes	Yes	Yes	Yes	Yes
Nrest NHS Hosp characteristics	Yes	Yes	Yes	Yes	Yes
MSOA Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
PCT area x Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	61,029	61,029	61,029	47,467	47,467
R-squared	0.193	0.193	0.194	0.168	0.168
Number of MSOA	6,781	6,781	6,781	6,781	6,781

Notes: *** denotes significance at 1%, ** at 5%, and * at 1% level. Observations are at the MSOA year level, and the sample includes all 6781 MSOAs in England. The dependent variable in all columns is the number of admissions for an NHS-funded elective hip replacement amoungst MSOA residents, age/sex standardized to the English population in 2002. ISP5 takes the value of one if an ISP conducting at least 5 NHS-funded hip replacements was the closest provider of NHS hip replacements in a given year. ISTC20 and AQP20 restrict the treatment variable to the appropriate type of ISP. Relative and absolute distances to the nearest ISP and NHS Acute hospital are measured in straight lines from the centriod of the MSOA to the provider post (zip) code. In columns (4) and (5), we include a quadratic term in relative and absolute distance respectively. Analysis is restricted to the period beyond 2004/05 in columns (4) and (5) as distance to large ISPs are undefined prior to 2004/05. Standard errors are robust to the presence of heteroskedasticity and clustered at the Primary Care Trust (PCT) area level.